Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
29 SEP 2003	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Methods to Account for Accelerated Semi-conductor Device Wearout in Longlife Aerospace Application	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **The Department of the Air Force, AFIT/CIA, Bldg 125, 2950 P Street, Wright Patterson AFB, OH 45433**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**128**	

ABSTRACT

Title of Dissertation: METHODS TO ACCOUNT FOR ACCELERATED SEMICONDUCTOR DEVICE WEAROUT IN LONGLIFE AEROSPACE APPLICATIONS

Joerg Dieter Walter, Doctor of Philosophy, 2003

Dissertation directed by: Professor Joseph B. Bernstein
Department of Mechanical Engineering

The aerospace industry is concerned that as semiconductor feature sizes are reduced future technology generations, device lifetime will decrease as well. Inherent device failure mechanisms, such as electromigration, hot carrier effects and time dependent dielectric (oxide) breakdown, may lead to shorter lifetimes at these smaller feature sizes. Many longlife aerospace applications must use commercially available off-the-shelf devices. The reliability margins in future devices may be decreased as semiconductor suppliers trade performance for reliability to meet the requirements of their core markets. If the lifetime of future devices proves to be inadequate for longlife aerospace applications, operating them at a derated stress condition can extend their lifetime. This is accomplished by reducing the operating voltage of the devices.

METHODS TO ACCOUNT FOR ACCELERATED SEMICONDUCTOR DEVICE WEAROUT IN LONGLIFE AEROSPACE APPLICATIONS

by

Joerg Dieter Walter

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2003

Advisory Committee:

Professor Joseph B. Bernstein, Chair
Professor John S. Baras
Professor Michel Cukier
Professor Mohammed Modarres
Professor Carol Smidts

ACKNOWLEDGMENTS

This work is sponsored by the Aerospace Vehicle Systems Institute (AVSI). The project principle investigator is Lloyd Condra of the Boeing Phantom Works. The principle investigator at the University of Maryland, and my academic adviser, is Dr. Joesph Bernstein. Other students on this project include Jin Qin, who analyzed our avionics field failure rate data and Bing Huang who is leading semiconductor device testing.

My education, and work on this project, was sponsored by the United States Air Force and the Air Force Institute of Technology.

The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The aerospace industry, including military, space and commercial users, has for the past several years faced a growing challenge in the use of commercial-off-the-shelf (COTS) semiconductor (Integrated Circuit (IC)) devices. This challenge has come from the electronic industry's concentration on the computer, networking, telecommunications and consumer markets and their resulting neglect of the aerospace market [1]. Continued developments within the semiconductor industry are further jeopardizing the ability of the aerospace industry to use future devices as they have used COTS devices in the past. In particular, shrinking device features pose the possibility of early wearout resulting in semiconductor devices which have a shorter expected lifetime then the systems in which they are incorporated.

The maturing of the semiconductor market over the last decades has shifted the emphasis of the industry away from the defense and aerospace markets to the commercial and consumer markets. Market pressures are driving a continued quest for more speed, larger memories and lower costs. This has been achieved by pushing device feature sizes down to less than 0.1 micron (nanometer scale). These advances

are widening the division between the requirements of the aerospace industry—harsher environments and longer expected lifetimes—and the needs of the consumer market. To gain increased performace along with lower cost, device manufactures have reduced or eliminated the reliability margins that allowed the devices to be used in aerospace applications. The expected lifetime of such devices in aerospace applications are being reduced from decades to years [1]. Increasing the problem for aerospace, the details of a manufacturer's processes, products and data are proprietary and are only shared with significant customers in target markets. This does not allow aerospace Original Equipment Manufactures (OEMS) to understand part capabilities or the risks of using the advanced technologies. Further obstacles to using COTS devices include the constant "improvement" of device designs, fabrication and assembly processes and test methods. These obstacles to using COTS devices maks it a challenge to take advantage of the advances in semiconductor technology and incorporate it into military and commercial avionics systems.

The answer of how to best incorporate COTS semiconductor devices into aerospace systems is a large question with technological, business and regulatory aspects. Today, several aerospace companies are working together with government, higher eductation, and semiconductor device suppliers to develop an Integrated Aerospace Parts Acquisition Strategy (IAPAS) [1]. This work supports the IAPAS effort and examines a small subset—early device wearout—of the overall problem. The next section of this introduction provides background information on the technological trends in the semiconductor industry, the business climate within the aerospace and how this effects

the aerospace industry. Following this is the problem statement along with an overview of the dissertation.

1.1 Background

1.1.1 Technological Trends

Technology trends are driven by a relentless quest for more processing power. Customers have come to expect a constant increase in power and functionality from IC devices while the cost remains constant or drops. To stay competitive in this market all device manufactures must constantly improve their products. In the past, the rate at which semiconductor products improved has followed Moore's law.

In 1965, Gorden Moore of Intel observed that the number of transistors in an IC grew at an exponential rate over time [2]. His "law" states that the number of transistors on a semiconductor device will double every 18 months. He expected this growth rate to continue into the future and history has proven him to be correct. Even today, the growth rate of transistors in future devices is expected to continue following Moore's Law [3].

The leading source in documenting how the semiconductor industry will meet this challenge is the *International Technology Roadmap for Semiconductors (ITRS)*. Published semiannually, this report is a cooperative effort of the world's top five region semiconductor associations, each association being comprised of the leading

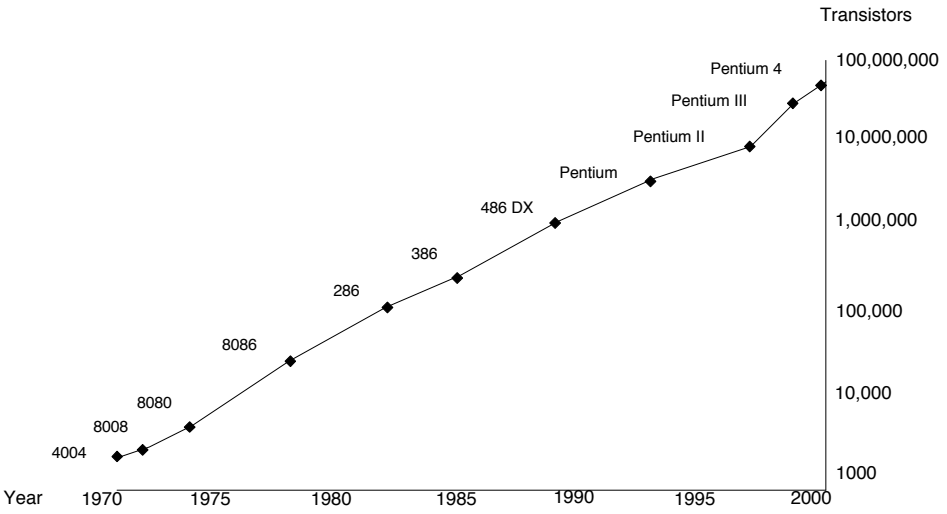

Figure 1.1: Intel Processor Growth. The growth in transistor count for Intel processors followed Moore's Law [3].

semiconductor device manufactures. It's purpose is to detail the technological directions of the semiconductor industry with the goal of continuing to meet the pace defind by Moore's law for the next fifteen years. The focus of the roadmap has been on tradition Complementary Metal-Oxide-Silicon (CMOS) circuits and highlights the technologies, trends and roadblocks of the technological advance.

The technological trends in the semiconductor industry are of critical importance in understanding the impact of early device wearout. As the number of transistors has grown, the typical die size of an IC has remained constant. This means the size of the device features on a IC die have shrunk dramatically in order to allow for more transistors to be placed in a device. These smaller device features, approaching a few

atoms in size, are causing the concern about lifetime. In the past, there has been adequate margin so that wearout was never an issue. But this margin may not be there in future devices, so it important to understand where semiconductor technology is headed in the future.

The principle method the semiconductor industry has used to meet the challenge of Moore's law has been the constant decrease in the minimum feature size of ICs [4]. This decrease in feature size is termed 'scaling'. While scaling has worked in the past, CMOS devices are now reaching the point (9 nm feature sizes by 2016) where it may be difficult to scale them further in the future. And it is at these small scales that problems of early wearout may appear.

1.1.2 Scaling

Scaling of semiconductors traditionally happens in a discreet fashion as manufactures move from one technology 'node' to another. A technology node is defined as the half-pitch of the smallest device feature printed. Half-pitch is defined as the spacing between device features (see Fig. 1.2). For DRAM (Dynamic Random Access Memory), the node is defined by half-pitch of the first-level metalization interconnect lines. For logic devices, such as microprocessors (MPU), the logic interconnect half-pitch refers to the first polysilicon or metalization layer. The half-pitch of MPUs and ASICs (Application Specific Integrated Circuits) lags behind the half-pitch employed for DRAM. Each node typically represents a decease in half-pitch of

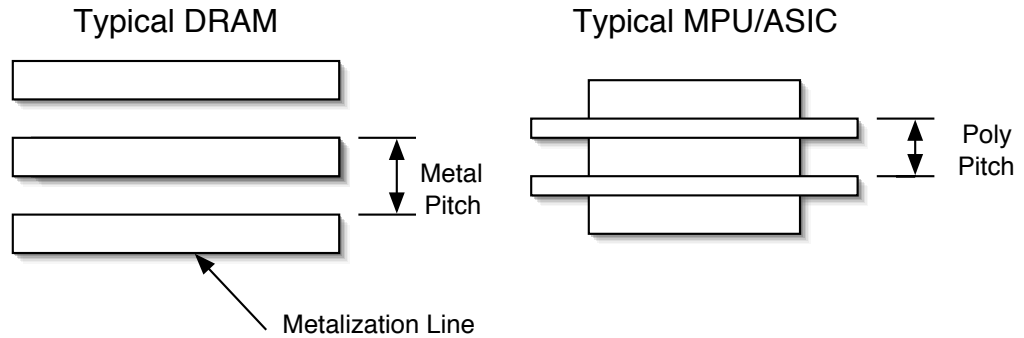

Figure 1.2: Half-Pitch. The pitch of a typical DRAM and Microprocessor (MPU)/ASIC. Half-pitch = (Pitch/2).

approximately 70% from the previous node or 50% from two technological levels back. This is illustrated in the Fig. 1.3.

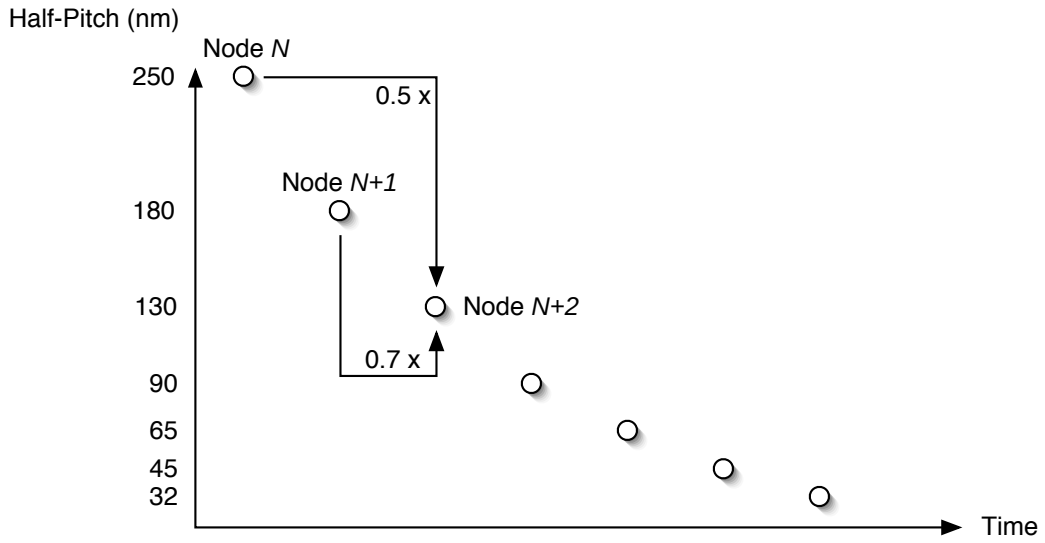

Figure 1.3: Half-Pitch Nodes. A schematic of the shrinkage of the half-pitch at processive technology nodes. With vertical axis represents the half-pitch of each technology node while the horizontal axis represents time.

According to The ITRS 2001 report [4], the rate of transition between technology nodes as accelerated from the traditional three years to two years between nodes for

microprocessors and ASICs. The rate is expected to continue until at least 2004 when a 90 nm half-pitch is reached. For DRAM, the time between nodes remains three years.

With technology nodes defined by half-pitch, the dimensions and parameters of other device specifications may be determined by applying scaling rules. One commonly used set of rules is that for constant field scaling. The principle of this rule is that device voltages and dimensions are scaled by the same factor (κ) such that the electric field ($\mathscr{E}$) remains constant. Scaled parameters are defined in Table 1.1.2. A diagram illustrating the use of scaling factors is shown in Fig. 1.4.

Table 1.1: Scaling Factors. Scaling parameters for MOSFET device parameters [5]

	Device Parameters	Multiplicative Factor ($\kappa > 1$)
Scaling Assumptions	Device dimensions (T_{ox}, L, W)	$1/\kappa$
	Voltage (V)	$1/\kappa$
Derived device parameters	Electric Field ($\mathscr{E}$)	1
	Depletion-layer width (W_d)	$1/\kappa$
	Capacitance ($C = \mathscr{E}A/t$)	$1/\kappa$
	Inversion-layer charge density (Q_i)	1
	Current,drift (I)	$1/\kappa$
	Channel resistance (R_{ch})	1

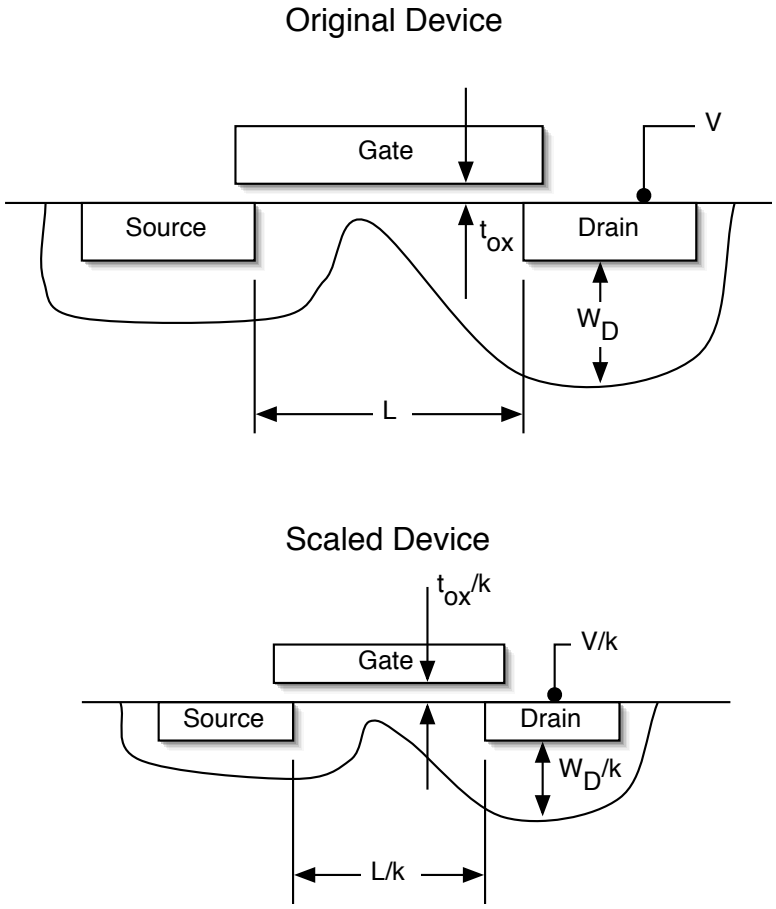

Figure 1.4: Scaling Factor. Application of the scaling factor to create a scaled device [5].

1.1.2.1 ITRS Roadmap Trends

The 2001 ITRS roadmap provides a more detailed set of predictions on scaling trends than can be derived using the scaling factors. As an example, the half-pitch trends for high power microprocessor devices is shown in Figure 1.5.

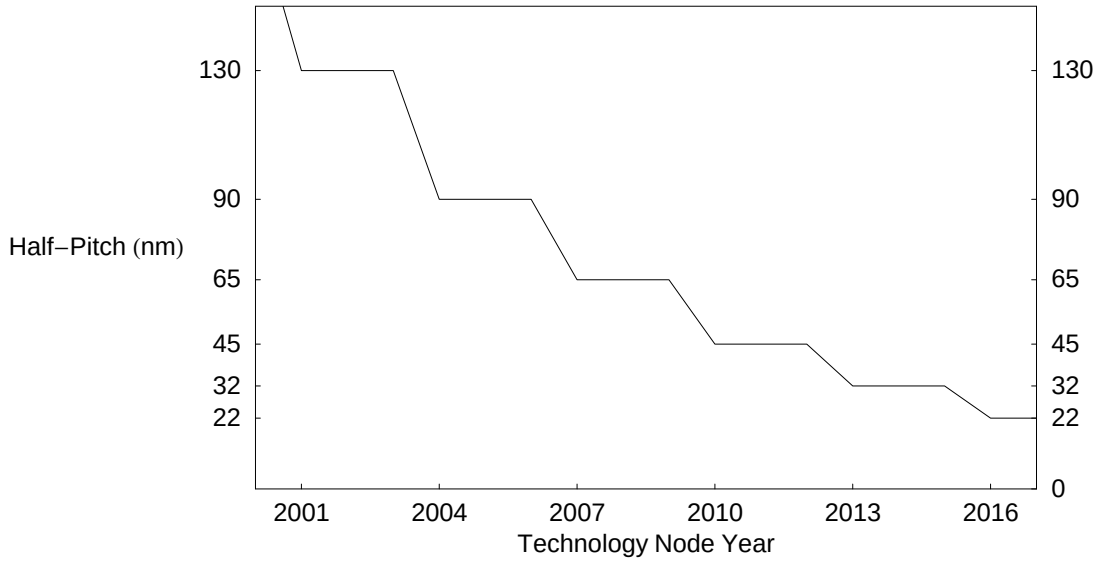

Figure 1.5: Node Size by Year. ITRS predicted node sizes for high power microprocessors.

1.1.2.2 Reliability

The ITRS report didn't provide a great level of information on reliability tends. It did however, state that each new technology requires new materials and techniques. This will introduce new failure regimes and defects. It leaves reliability as an area of concern which places challenges on testing and wafer level reliability (WLR).

Difficulties listed include the reliability of very thin oxy-nitride gate dielectrics due to high gate leakage, new gate electrode materials, non-classical CMOS structures and the reliability of high κ gate dielectrics.

1.1.3 Aerospace Industry Business Climate

From a technological point of view, there is no problem with using older qualified designs and technology in aerospace systems. The problem in the aerospace industry stems from the business environment. Aerospace companies are currently faced with a swiftly diminishing set of manufacturing sources [1] as the electronics industry focuses its efforts on computer, networking, telecommunications and consumer products. The military/aerospace markets have shrunk as a percentage of the overall semiconductor market and do not carry the clout they once enjoyed. In addition, the Secretary of Defense Perry Acquisition Reform Memorandum (Perry Memo) [6] of 1994 strongly encouraged the military services and defense contractors to eliminate the use of military specifications and standards.

By eliminating the need for new mil-spec components for military applications, the Perry memo resulted in a decline in the availability of these components for the entire aerospace industry. The commercial side of the aerospace market was the first to transition to COTS parts. Military products are transitioning at a slower pace [1]. Initially, the fear was these parts would not be suitable for harsh aerospace environments. However, it was shown that improvements in COTS parts had made

them acceptable for aerospace use in all but the most demanding applications (such as radiation sensitive space applications).

While COTS semiconductors have proven themselves in service, there is no guarantee they will remain up to the challenge of military/aerospace applications in the future. To compete in their core markets, semiconductor device manufactures must constantly improve the price-performace ratio of their products. This is accomplished by scaling device features down and increasing the transistor count in accordance with Moore's law. As device feature dimensions (half-pitch, line widths and gate lengths) shrink to less than 0.1 microns, there is potential for impacts on aerospace users. These impacts include [1]

- Service lives of 3-10 years for aerospace applications.
- Possible non-constant failure rates.
- Increased susceptibility to atmospheric radiation.
- Changes in configuration due to constant "improvements" to device design and manufacture.
- The inability of aerospace users to understand the impact of nanometer technology used in aerospace applications.

If the aerospace industries concerns are confirmed, decreased IC reliability will have a direct impact on the supportability of aerospace systems. Shorter component lifetimes, or early wearout, would drive increased maintenance costs as failed

components and Line Replaceable Units (LRU) must be replaced. For example, one estimate, assuming a six year lifetime on semiconductor devices, shows that if LRUs had to be replaced when the component devices failed, the added support cost for a commercial aircraft could reach one million dollars a year[1]. Additional costs would include the need for redesign and certification since components in production when the original boxes were built are unlikely to remain in production over the lifespan of the aircraft. As the LRUs need to be replaced, the replacement boxes will have to use the newer technology and devices available at that time.

With these unknowns and concerns, several aerospace companies, such as Boeing and Honeywell, are beginning the process of formulating an approach to dealing with these issues. One aspect of this is to support research into the subject. This consists of several different tracks. First, Boeing, Honeywell and the Defense Standardization Program Office are working on a program to develop an Integrated Aerospace Parts Acquisition Strategy (IAPAS). Part of this program's efforts include colaberative research with the Aerospace Vehicle Systems Institute (AVSI). The purpose of these activities include understanding and addressing the impacts of nanometer technology. One of the projects is AVSI Project #17 — *Methods to Account for Accelerated Semiconductor Device Wearout*.

[1]This assumes 300 avionics boxes (LRU) per aircraft, a $20,000 cost per box and replacement of each box every six years.

1.2 AVSI Project #17

The purpose of AVSI project #17 is understand the impact of shrinking device features and its implications on device lifetime and the potential for device wearout. The premise behind the project is that future avionics systems must be designed, produced, operated, maintained, and supported using COTS components. But with trends in the semiconductor industry moving counter to aerospace needs, the aerospace industry cannot assume that the design, production or service life of individual LRUs will be greater than 5-10 years [7]. Specifically, the three areas of interest for this project are the inherent device failure mechanisms of electromigration, hot carrier effects and time dependent dielectric (oxide) breakdown (TDDB).

1.2.1 Work Packages

AVSI Project #17 has a total of eight Work Package/Milestones (WP/MD). The responsible parties include the AVSI members sponsoring the project and the University of Maryland. The project was approved and formally started Spring 2002. My research addressed three of the work packages these were:

1. *Determine Likely Failure Mechanisms of Future Semiconductor Devices in Avionics Applications*. Conduct a literature search and consult with semiconductor device manufacturers to determine likely failure mechanisms of future semiconductor devices in avionics applications. Obtain design information regarding expected device lifetimes from device manufacturers. The deliverable

from this WP/MD is a report with quantitative information regarding the above topics.

2. *Develop Models to Estimate Expected Lifetimes of Future Avionics*. Based on published information and roadmap information regarding future avionics designs, develop mathematical models to describe time-to-failure of future semiconductor devices in aerospace applications. This will involve making some assumptions and customizing the models to fit aerospace conditions. The deliverable is a report with equations and sample calculations to estimate time-to-failure with respect to the failure mechanisms identified in WP/MD #1.

3. *Develop Device Assessment Methods and Avionics System Design Guidelines*. Using the information developed in the previous WP/MD, develop guidelines and, if necessary, suggest test methods to evaluate the potential lifetimes of specific semiconductor devices in existing and future avionics systems. Also, develop design guidelines for future avionics systems to minimize effects of early device wearout.

1.2.2 Project Scope

While the WP/MS define the goals and direction of the research, they didn't explicitly define the scope of the work. AVSI Project #17 is only a part of the Integrated Aerospace Parts Acquisition Strategy. The scope of this research is on understanding the inherent failure mechanisms of semiconductor devices that could lead to early

wearout or a reduction in life time. Additionally, this project includes developing ideas and methods to minimize the impact of any wearout potential or reduced lifetime. Excluded from this project are extrinsic failure mechanisms such as radiation, packaging and electrostatic shock.

1.3 Report Overview

After this introduction, the next chapter (Ch. 2) discusses the systems engineering methodology used on this project. The following two chapters will detail the research. The first (Ch. 3) serves as a tutorial, discussing the failure mechanisms of semiconductors and how scaling will effect these mechanisms. The next chapter (Ch. 4) discusses derating semiconductor devices to increase their lifetime and reliability. The last chapter (Ch. 5) summarizes the results to date and discusses future work necessary to continue supporting this project.

Chapter 2

Methodology: A Systems Engineering Process

Upon an initial look, the problem of AVSI Project #17, understanding the impact of shrinking semiconductor device features and its impact on device lifetime, appears to be a reliability problem. If the root causes of the failure mechanisms and their relationships with shrinking device features can be understood, then design, process or manufacturing changes can be made to alleviate the problem. However, in actuality AVSI Project #17 is much more than just a technological problem. It is a system level problem with technological, engineering, business and market aspects and it involves a wide range of actors.

Because of this, it was approriate to incorporate some Systems Engineering methodologies into the research process. This chapter reviews several Systems Engineering methodologies. Next the chapter explains why it is important to incorporate Systems Engineering methodologies into the project and how they are implemented.

2.1 What is Systems Engineering?

An simple, agreed upon explaination what Systems Engineering is doesn't exist. The best place to start in understanding Systems Engineering is with the definition given by *The International Council on Systems Engineering* (INCOSE) [8],

> "Systems Engineering is an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, then proceeding with design synthesis and system validation while considering the complete problem: Operations, Performance, Test, Manufacturing, Cost & Schedule,Training & Support, Disposal. Systems Engineering integrates all the disciplines and specialty groups into a team effort forming a structured development process that proceeds from concept to production to operation. Systems Engineering considers both the business and the technical needs of all customers with the goal of providing a quality product that meets the user needs."

A more concise definition is given by Austin [9]: Systems Engineering is "the end-to-end development—planning, analysis and design, implementation, operation, retirement—of complex engineering systems, taking into account engineering and business concerns". Both these definitions mean the same thing: Systems Engineering takes into account all parts of a complex system, including both engineering and non-engineering aspects.

Decreasing lifetime of semiconductor devices is a Systems Engineering problem because aerospace companies do not have the purchasing clout to strongly influence the design of advanced ICs. This leaves them in the position of having to purchase those devices which are brought to the market and intended for other types of applications. If this were solely a reliability problem, then the design or manufacturing process of the IC devices would be modified to increase their lifetime, alleviating the problem. Since this isn't an practical solution, the remainder of the system must be altered to accommodate the reliability shortfall.

As a Systems Engineering problem, all aspects of an aerospace system, including operations and maintenance, are subject to consideration. Systems Engineering depends on a systematic process for problem solving. Many Systems Engineering tools and methodologies have been proposed. The exact tool used depends on the nature of the task.

2.2 Systems Engineering Methodologies

One of the earliest System Engineering methodologies was proposed by Hall [10]. This methodology serves as the basics of many later ideas on applying Systems Engineering principles. Hall's method is an iterative model with each iteration divided into seven steps. Each iteration serves to improve the definition of the system concept and design. Hall's seven steps are:

- Problem Definition: Explicitly define the problem at hand along with constraints

and the scope.

- Value System Design: Define a system to quantitatively or qualitatively score alternative system designs and concepts against.
- Systems Synthesis: Create a set of system designs and concepts.
- System Analysis: Define the systems.
- Modeling and Optimization: Model and refine the system concepts.
- Decision Making: Select an alternative, or set of alternatives, to bring into the next iteration.
- Planning for Action: Plan the course of action for the next iteration or for implementation.

These steps are not firm fixed rules and many authors have proposed additions, modifications and clarifications to this basic model. One particular difficulty with Halls model is that the three steps, Systems Synthesis, System Analysis and Modeling and Optimization, are not distinct. They tend to be parts of the same whole. Other authors have refined Hall's basic model. In particular, I examined the 'Model-Based Method' [11] derived from Hall's model with improvements taken from papers by Sage, Hill and Warfield, and Mosard [12, 13, 14, 15].

The 'Model-Based Method' consists of five steps. These are:

- Problem Definition

- Model Definition
- Modeling and Analysis
- Decision Making
- Implementation

The last methodology I examined was developed by the Institute for Systems Research (ISR) at the University of Maryland [9]. This methodology is a visual modeling language, based on Unified Markup Language (UML) diagram notation, for systems architecting and engineering design. While it doesn't appear to be directly descended from Hall's, the ISR methodology does share many of the same inherent features as the other methodologies I reviewed, just with different ways of accomplishing them. Two aspects of the ISR approach that I favored as an improvement were its emphasis on traceability and on the use of visual UML modeling.

Traceability refers the process of ensuring each function and feature can be traced back to the original requirement, either directly or as a derived requirement. Tracability works two ways. First, it is used to ensure that all the requirements are implement in the design solution. Secondly, it ensures that the design solution doesn't include extraneous features unnecessary to meet the requirements. At this point in AVSI Project #17, traceability isn't a critical tool, but as the development of a solution progresses it will become increasingly important.

The second feature of ISR's methodology I liked was the use of UML modeling. UML is an object-oriented visual modeling language typically used in software

development. The language is extensible allowing it to be adapted to modeling physical as well as logical systems.

2.2.1 Project Methodology

In examining the system methodology models I found them to be focused on design solutions to a problem. But for this project I needed a higher level of abstraction for the starting point. The first phase of this project is to understand the nature and implications of the problem, not to necessarily design a fix. So the methodology required would have to accommodate this search of the problem space.

The methodologies drawn on so far are iterative. They are best suited for developing and examining alternative solutions to a problem and then refining in them until a solution is reached. Each iteration serves to narrow the solution space (number of alternatives) and to increase the level of detail of the remaining solutions. At the conclusion of the process and solution is implemented.

These methodologies were not directly suited to the AVSI Project #17 project needs at this point. This project is a research project with the true problem and needs unknown. This project was not well defined, as in a purpose the project was to define itself. A spiral development model is best suited for this since it allows knowledge gained to be brought back into the research cycle. The purpose of each spiral is to refine the understanding of the problem and to propose solutions and/or directions for continued work. The iterative methods provided structure for the spirals. I based each

spiral on the ‘Model-Based Method’, but consolodated the Decision Making and Implementation steps.

As in Hall’s method, the first step of the spiral model is to define the problem. This involves writing out the statement of what the problem is and it defines the system boundaries and the scope of the problem.

The second step is model definition. This step involves building a model of the system. This step is where UML modeling may be applied to good advantage. Another portion of this step involves gathering information to flesh out the model.

The third step, Modeling and Analysis, focuses on taking all the model inputs and understanding their impacts on the system.

The forth and last step is Decision Making. Here the results of the previous steps are used to draw conclusions and make decisions on the future direction of the research. The knowledge gained in these four steps in fed back into the spiral development pattern and it continues again.

With this model, it is possible to call a halt to the spiral if conditions warrant. An example of this would be when the state of knowledge in the system is such that a suitable level of detail has been reached in the problem defintion and understanding that a solution can be developed. At this point the methodology would continue to the implementation step and then enter a pattern of design iterations to refine the solution. It is also possible for design solutions to be broken out of the spiral model to be developed on their own in a seperate iterative process as the main spiral explores other alternatives. The next two subsections describe of how a Systems Engineering

methodology was implemented in this research.

2.2.2 First Spiral: Device Physics

The first action in this spiral is to define the problem. The overarching problem definition was provided by AVSI Project #17 and states:

> "This project will develop methods to evaluate the mechanisms and accommodate the effects of accelerated semiconductor device wear out on avionics system design, production, and support; and develop methods to account for shorter device lifetimes in avionics system safety and reliability analysis" [16].

Model definition begins with a use case model. The initial use case model, shown in Figure 2.1, provides a picture of the actors involved with the system. Its purpose is to highlight the different actors having an effect (extends) on the lifetime of semiconductor devices, as well as the actors being impacted (uses) by the device lifetime. Highlighting the different actors involved with the system demonstrates how this is a Systems Engineering problem and the entire scope of the problem must be considered when understanding the probelm and synthesizing solutions.

The next step in Model Definition is researching and understanding failure in semiconductor devices. This includes determining the suitable lifetime models for the three wearout failure mechanisms (electromigration, hot carrier effects and TDDB). The Modeling and Analysis step involves understanding how the mechanism lifetime

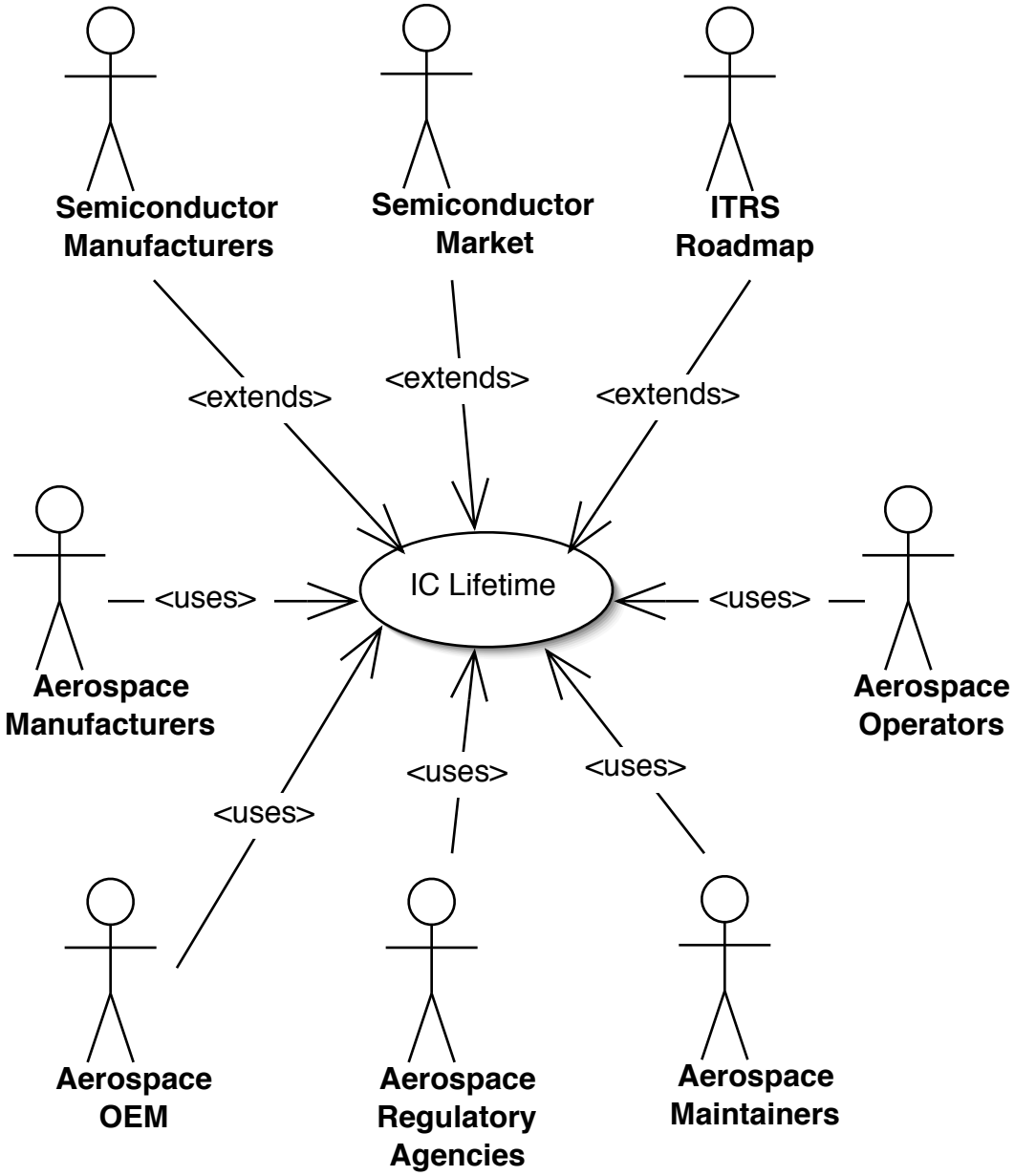

Figure 2.1: Use Case Model.

models react to variations in their input parameters as well as the effect of device scaling on the lifetime for these mechanisms. The Decision Making and Implementation steps involve moving to the next spiral, derating the semiconductor device for increased lifetime. The results of this step are presented in Chapter 3.

2.2.3 Second Spiral: Derating

The second spiral focuses on exploring the concept to alleviate the impact of shrinking device lifetime, derating the semiconductor devices for aerospace use. The problem definition for this spiral is:

> *Model the change in semiconductor device lifetime, for a device operated a derated conditions, from electromigration, hot carrier and TDDB failure mechanisms. The models shall be usable by AVSI Project #17 members, with data available to them (either from the device manufactures or via accelerated life testing) for the purpose of estimating lifetime improvement from device derating.*

This Model Definition step draws on the results of the first spiral to define a dearating factor for each of the failure mechanisms. The Modeling and Analysis step involves verification of a constant failure rate assumption for each of the mechanisms and an analysis of the derating models response to changes in the input variables. The results of this step is presented in Chapter 4. The results of the Decision Making and Implementation steps are covered in Section 5.2, Future Work.

2.3 Summary

This chapter has been a brief overview of the methodology used on this project. At this early stage in the AVSI Project #17, the most significant contribution of the Systems Engineering approach was to consider the impact of non-technical aspects on the semiconductor device lifetime and how that will effect solutions to the problem. Future work, such as developing specific design solutions using derated devices, will require a more rigorous iterative Systems Engineering process.

Chapter 3

Background: Impact of Scaling

This chapter examines the effect of device scaling on the inherent reliability of semiconductor devices. It begins with basic concepts of failure and the classification of failure. Next, the mechanisms of electromigration, hot carrier degradation and oxide breakdown are explained and models are presented for predicting mechanism reliability. At the end of this chapter, the impact of technology node scaling on lifetime is discussed.

3.1 Understanding Failure

3.1.1 What is Failure, Degradation and Wear-out

A semiconductor device has *failed* when response parameters from the device (e.g. voltage, capacitance, resistance, gain, etc.) no longer meet the design parameters. A simpler way of stating this is the device has failed when it is in a physical state or condition in which it can no longer perform its intended function. The failure may have been caused by a sudden internal or external event triggering a physical change in the

device or the failure may have developed slowly as physical changes over time alter the response of the device. The latter is referred to as *degradation*. *Wear-out* occurs when degradation reaches the point where the device is considered to have failed.

A failure may be classified into one of three categories, *intrinsic*, *extrinsic*, or *electrical stress* (in-circuit) failures. Of these, the areas of greatest intrest in this study are the intrinsic failure mechanisms.

Intrinsic failures are the result of failure mechanisms originating with the semiconductor device, or die, and the processing during the ‘front end’ of manufacturing. Examples of these mechanisms include design errors, lithography and processing defects, contamination or the limitations of material properties. These defects may result in a device being fatally defective so it never functions or these defects may be small enough so they are non-lethal. However, stresses from temperature, voltage and current flow, along with humidity and radiation, may cause these non-lethal defects to grow into a lethal defect resulting in a failure.

Extrinsic failures are identified with the interconnection and packaging of chips in the ‘back end’ of manufacturing. These types of failures are external to the device circuitry itself. Extrinsic failures are not the subject of the this research since they are not directly related to the device itself or shrinking device features.

Electrical stress failures are generally caused by discrete events and are often considered to be ‘random’ failures. These damaging events typically occur during handling and they include Electrostatic Discharge (ESD) and Electrical Overstress (EOS). In fact EOS and ESD can make-up over 50% of in field failures [17]. EOS is

caused by events which occur during normal circuit operation and lead to over-voltage and over-current stresses of long duration—greater then 1 ms, although some stress events as low as a few μs long are classified as EOS (and may be too fast for protection schemes to prevent). The effect of EOS is typically to cause a hot spot to develop in the IC. As it gets hotter, more current flows into the heated region and temperatures continue to build. When the temperature approaches the melting point of Si (1688 K), failures may occur as short circuits form in junctions or metallization melts creating open circuits [17]. ESD is caused by stress extrinsic to the normal operation of a device. An example is static charges of over 100 V. If an IC isn't protected, these charges may damage the gate oxides in Metal-on-Silicon (MOS) transistors. Protection is provided in nearly all circuits, so the typical IC failure mechanism is thermal [18].

All failure types may occur at any point in a semiconductor device's lifetime. The 'Bathtub' curve, shown in Figure 3.1, represents the instantaneous failure rate at any point over a device's lifetime. This is a hazard rate and is measured as the number of failures per unit time[19]. The curve is divided into three phases, infant mortality, useful life and wearout.

For a large sample of devices, the infant mortality phase represents items that fail early due to manufacturing defects. As defective items quickly fail and are removed from the system, the failure rate drops. Burn-in is often used to screen out items inclined to fail prematurely, resulting in the remaining devices having a longer expected lifetime. However, burn-in is only justified so long as the failure rate decreases over time as seen in the infant mortality region of Figure 3.1.

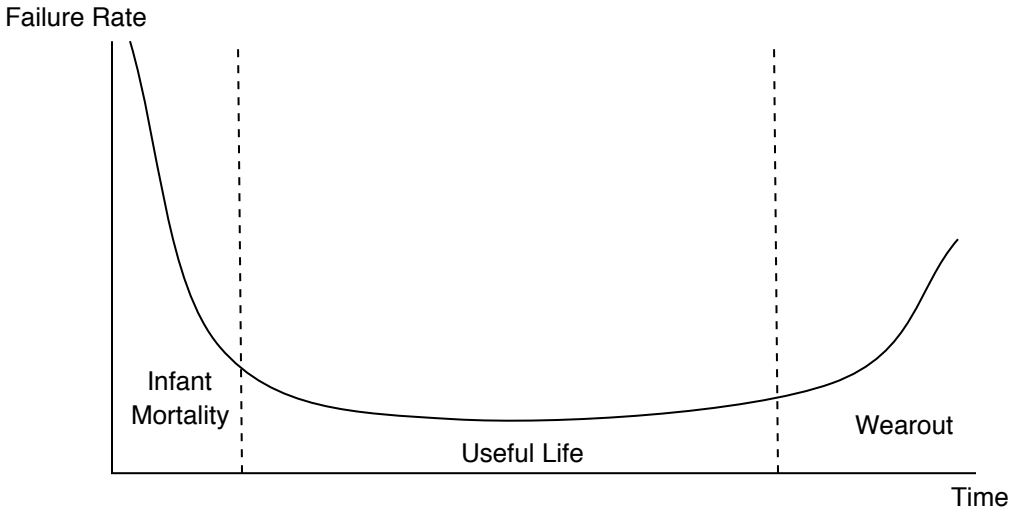

Figure 3.1: Bathtub curve. The line represents the instantaneous failure rate of a device showing the periods of infant mortality, useful life (random failure) and wearout.

For items that survive the infant mortality phase, they enter a long period of a relatively constant failure rate. This period of time is termed the useful life of the product. Failures during this period are attributed to random failures resulting from sources such as random external events, non-lethal intrinsic defects that have grown into lethal defects, or early wearout. The end of the product's useful life comes as the failure rate begins to climb.

An increasing failure rate represents the start of the wearout phase. During this phase the physical degradation suffered throughout the working life of the item begin to increase the failure rate.

3.1.1.1 Infant Mortality Failures

Failures during the infant mortality period are typically the result of defects introduced during the manufacturing process. Infant mortality need not be a direct concern of an IC user—assuming the manufacturer has a burn-in procedure to screen out susceptible parts.

The manufacture of semiconductor devices will always result in products with defects. Some of these defects are insignificant, others serve as a nucleus of failure in the future, while still others are significant enough to render the device useless or to cause it to fail quickly. If the infant mortality failure rate distribution of a product is known, infant mortality by itself should not directly pose a reliability problem for devices placed in service. Screening, through burn-in, weeds out the devices that have significant manufacturing defects. In a burn-in process, each manufactured device is tested at normal or accelerated stress conditions. The length of the burn-in period is set to correspond to the length of the infant mortality period. Devices that survive the burn-in are expected to be susceptible only to random and wearout failures. Through the use of conditional probability, it can be shown that for a product with a decreasing failure rate, the expected life of surviving product after burn-in is greater than if there had been no burn-in [20].

Infant mortality is important to understand because it is an indicator of the overall reliability of a device in that defects that don’t cause initial failures often serve as sources of failure in the future. Hence a product experiencing high infant mortality

usually indicates a product with the potential for a higher failure rate during its useful life. Additionally, high infant mortality may be indicative of inadequately controlled manufacturing processes or design parameters that are not optimal. Likewise, yield[1] in IC production has been shown to relate to in-service reliability [18].

3.1.2 Yield

During the manufacturing process, there are several potentially lethal defects that may occur, leading to reduced yield or infant mortality failures. Among these are processing errors, contamination, material flaws, residual stresses and contact failures. Yield refers to the percentage of devices manufactured that are suitable for sale to a customer.

Contamination of wafers during fabrication is an ever present source of defects in semiconductors. As the process size has decreased, the size of particles that may cause defects have shrunk. Clean rooms have reached fantastic levels of cleanliness, but now individual molecules, such as 0_2, N_2, & H_2O, are sources of contamination. Etching and cleaning liquids and wear particles from hardware contribute to the contamination problem as well. The greatest impact of contamination is to yield, but contamination creates non-lethal defects as well. Examples of contamination sources include ion implantation, metal contamination and damage to lithography masks.

Fabrication also introduces stress into the ICs. While it isn't a defect by itself, stress may lead to the formation of defects. Stress results from the processing of

[1] In semiconductor fabrication, yield is the percentage of functional devices manufactured versus the total number manufactured.

dissimilar materials in contact with each other. High temperatures are used during the manufacturing process and as every material has different thermal expansion properties, they expand and contract differently during fabrication, leading to the formation of internal stresses. Stress may also be introduced during grinding operations to thin the backside of wafers. Defects occur when excessive tensile stresses cause cracks to form in films or when compressive stresses cause wrinkling and the loss of adhesion between a film and substrate.

Defects are not only created during the fabrication process, but they may also reside in the material as well. The silicon wafers used in manufacturing will contain intrinsic defects [18] such as lattice vacancies, dislocations and grain boundaries. These defects may have a catastrophic effect on the electrical properties of device features, but are generally not a source of wearout, rather, intrinsic defects in circuits themselves.

3.1.2.1 Non-lethal Defects

Yield is a good indicator of the presence of non-lethal defects in a device. A consequence of non-lethal defects in an IC is that they result in some device features lower in strength than average. Low yield products typically have a lower reliability in service because of existence of an increased number of defects [18]. Examples of these defects ‘near opens’ and ‘near shorts’ in the metallization and near shorts in gate oxides and passivating insulators. Defects on a wafer, both lethal and non-lethal, are random distributed (baring a systematic source of flaws). The size, or criticality, of the defect will be random based on the location of the contamination or flaw. The result is

non-lethal defects will grow to failure after various periods of use, appearing to a user as random failures.

3.2 Fundamental Failure Processes

To move from an operating to a failed state, the physical state of a semiconductor must change. Without moving parts, the change in the physical change of a semiconductor results from the movement of electrons within it along with the associated electric fields, currents and temperatures. Two of the concepts necessary for understanding these physical changes are the Arrhenius model and the concept of diffusion.

3.2.1 The Arrhenius Model

The Arrhenius model forms the basis of understanding the lifetime prediction for many semiconductor failure mechanisms. It has generally been accepted that the best way to accelerate failure in electronics is by raising the temperature. While heat by itself is not a mechanism of failure, thermal energy does contribute to the acceleration of many failure mechanisms. Because some failure mechanisms are a thermally activated process, time to failure is often modeled through the use of a temperature dependent relationship. The standard way to do this is through the Arrhenius model. This model was originally developed in 1899 to model the reaction rate of chemical constituents. Over the years it has been used to model temperature acceleration factors for electronic component failure. The basic form of the Arrhenius model for predicting MTTF in

electronic devices is

$$MTTF = A \exp\left(\frac{E_a}{kT}\right) \tag{3.1}$$

where A is the frequency factor, E_a is the activation energy, k is Boltzmann's constant (8.62×10^{-5} eV/K) and T is absolute temperature (K). The activation energy may be described as an energy barrier separating the reactants from products in a chemical or physical process connected to a particular failure mechanism [21]. Often the activation energy of different failure mechanisms is discussed in literature. In this context, activation refers back to the Arrhenius equation. The higher the activation energy, the quicker the failure will occur with increased temperature. Most failure processes have a positive activation energy. One exception is *hot carrier degradation* (see section 3.4.2). However, recently it has been found that oxide related degradation is not accelerated by this model.

3.2.1.1 Acceleration Factor

A common application of the Arrhenius model is in the development of Acceleration Factors (A_f) in conjunction with accelerated life testing. During accelerated life testing, the lifetime of a device is experimentally determined at a high stress condition (such as increased temperature, voltage, current, etc.). An Acceleration Factor is used to extrapolate those results to the stress level at normal operating conditions in order to predict the lifetime of the device.

The definition of an Acceleration Factor is the "ratio of the measured failure rate of semiconductor devices at one stress condition to the measured failure rate of identical

devices stressed at another condition" [22]. The Mean Time to Failure (MTTF) of a device operating under accelerated stress conditions, multiplied by A_f, provides an estimate of the device's MTTF under normal use conditions. This defines A_f as

$$A_f = \frac{\text{MTTF}_{Use}}{\text{MTTF}_{Stress}} \tag{3.2}$$

3.2.2 Diffusion

Diffusion is the process whereby particles move from areas of a higher concentration to areas of lower concentration. Diffusion is a fundamental failure mechanism in that contaminant atoms may change the electrical characteristics of an IC. Semiconductor devices function because the silicon, or other semiconductor material, is doped with other elements to change its local electrical properties in order to build gates and other electrical components. For example, doping levels determine the properties of *n* and *p* junctions.

Diffusion of atoms from, or into, these doped regions will change the characteristics. When the characteristics are altered too much a failure is considered to have occurred. Diffusion plays a role in other failure mechanisms as well. For example, electromigration, corrosion and nearly all thermally activated processes are due to diffusion mechanisms.

The rate of diffusion is dependant on three factors: temperature, diffusivity of the migrating atoms and concentration of the migrating atoms. By examining the physics of diffusion we can see how these factors apply. An explanation of the physics begins

with Fick's law [18],

$$J_m = -D\frac{dC}{dx} \tag{3.3}$$

Fick's law defines the diffusion process where the flux in the positive x direction is given by J_m, D is the diffusion coefficient, and dC/dx is the concentration gradient. The diffusion coefficient is dependant on a number of factors. These include the nature of the diffusing atoms, the matrix, the transport method (lattice, grain boundary, dislocation, surface, interstitial, etc.), temperature and the concentration of the diffusing species. The relation of D to temperature is given by an Arrhenius type equation

$$D = D_o \exp\left(\frac{-E_D}{RT}\right) \tag{3.4}$$

with R being the gas constant and E_D the activation energy for diffusion.

The equation for non-steady state diffusion in one dimension is

$$\frac{\partial C(x,t)}{\partial t} = D\frac{\partial^2 C(x,t)}{\partial x^2} \tag{3.5}$$

when D is constant. This equation may be solved using standard partial differential equation techniques. The first step is to define the initial conditions by assuming a semi-infinite matrix with an initial concentration of C_o atoms. More atoms enter the matrix at $x = 0$ from a concentration of C_s. Next we assign the boundary conditions: $C(x,0) = C_o$, $C(0,t) = C_s$ and $C(\infty,t) = C_o$. The solution is

$$\begin{aligned} \frac{C(x,t) - C_o}{C_s - C_o} &= \text{Erfc}\left(\frac{x}{(4Dt)^{1/2}}\right) \\ &= 1 - \frac{2}{\pi^{1/2}} \int_0^{\frac{x}{2(Dt)^{1/2}}} \text{e}^{-z^2} dz \end{aligned} \tag{3.6}$$

If an instantaneous surface concentration (S) is assumed rather then a constant (C_s), a simpler solution can be found. This is a gaussian solution,

$$C(x,t) = \frac{S}{(\pi Dt)^{1/2}} \exp\left(-\frac{x^2}{4Dt}\right) \tag{3.7}$$

The extent of diffusion penetration at any given time may be estimated using

$$x^2 = 4Dt \tag{3.8}$$

3.3 Modeling Failure

The failure of semiconductors may be modeled in one of three ways, the physics of failure, through simulation or via statistical models. All three methods have limits and weaknesses.

Predicting the failure of a device from first physical principles is intellectually appealing. If you could accomplish this, you would have a complete picture of how a device would fail. This would allow you to precisely trade-off design and performance specifications for the necessary reliability. However, as a practical matter, a quick review of the literature about any given failure mechanism will show that there is usually insufficient knowledge to completely model the physics. This uncertainty drives the failure prediction from being a deterministic problem to a probalistic one.

Simulation, either discrete event or continuous, allows probabalistic effects to be incorporated into a physics of failure model. Where detailed knowledge of the physics is unknown, probalistic models of the effects may be substituted. Simulation is

certainly possible for device manufacturers who have detailed knowledge about their devices. This detailed knowledge is typically proprietary information, meaning other parties will have less information to work with, making simulation more difficult. Another weakness of simulations is the need to fully understand and model all interactions. This can easily result in complex models. According to Huntington [23], "Either one takes all the variables into consideration, then the problem is most likely not solvable, or one restricts the consideration to a manageable degree, in which case the model is probably not accurate."

Without detailed knowledge of the device physics, the best way to represent device failures is through probalistic statistical models. These models use observed relationships between failure times and various input parameters to generate probabilistic assessments of when failure may occur.

3.3.1 Reliability Distributions

The probalistic lifetime of semiconductor devices are quantified using a variety of terms. The most basic of these is the failure rate (λ) defined as the number of failures per unit time. This corresponds to the failure, or hazard, rate shown in the bathtub curve. Assuming a constant failure rate (as seen during the useful life phase), the reciprocal of λ is the Mean Time Between Failure (MTBF $= \frac{1}{\lambda}$). Since the failure rates are not always constant, a more generalized term, Mean Time to Failure (MTTF), is also used. The standard method of quantifying semiconductor device reliability is via

FIT (Failure in Time) rates. FIT rates are defined as the number of failures per 10^9 device-hours, e.g. 1 FIT = 1 failure in 10^9 device-hours [24, 25].

The reliability of a semiconductor is only partially defined by its lifetime. To know the probability that a semiconductor device is functioning at a given point in time requires a failure rate distribution. For this purpose, two of the most popular distributions used are the *Weibull* and *exponential*. The Weibull distribution is used because of its ability to assume a wide variety of shapes, including decreasing, increasing and constant failure rate models. For a constant failure rate, the Weibull distribution reduces to the exponential model.

The exponential model has many useful properties. First, it is memoryless, meaning the probability of failure at any given point in time is independent of how long the device has been operating previously. This makes the exponential a favored distribution because of its ease of calculation since it uses only a single parameter to describe the distribution. Additionally, the exponential distribution is often a good model for the failure of systems with a large number of components. Within a system, each component's failure occurs randomly, according to that component's various failure modes and expected life distribution. Given a sufficiently large set of components and various failure modes, the times of the individual random component failures average out to a constant system failure rate [26].

3.4 Wear-out Mechanisms

Semiconductor device wearout concerns intrinsic mechanisms, electromigration, oxide breakdown (TDDB) and hot carrier effects (HC) effects encompassing hot carrier degradation (HCD) and negative bias temperature instability (NBTI).

3.4.1 Electromigration

As electrons pass through a electrical conductor they will transfer some of their momentum to the conductor's atoms. If the current density is high enough some of those atoms will be pushed along by the electron flow, depleting material at the cathode side and building up excess material at the anode. This diffusive process is known as *electromigration* and results in failure though damage from the formation of open circuits, increased electrical resistance or short circuits.

Failures from electromigration occur as voids and/or hillocks form within the semiconductor device. Short circuits are the result of hillocks breaking the oxide layer, allowing the conductor to come in contact with other device features. Alternatively, voids and microcracks may increase the resistance in a conductor as the cross sectional area is reduced. While the increased resistance alone may result in device failure, the increase in local current density and temperature, resulting from the increased resistance, can lead to thermal runaway and catastrophic failure [27], such as an open circuit failure. Other types of damage include whiskers, thinning, localized heating, and cracking of the passivating dielectrics [18].

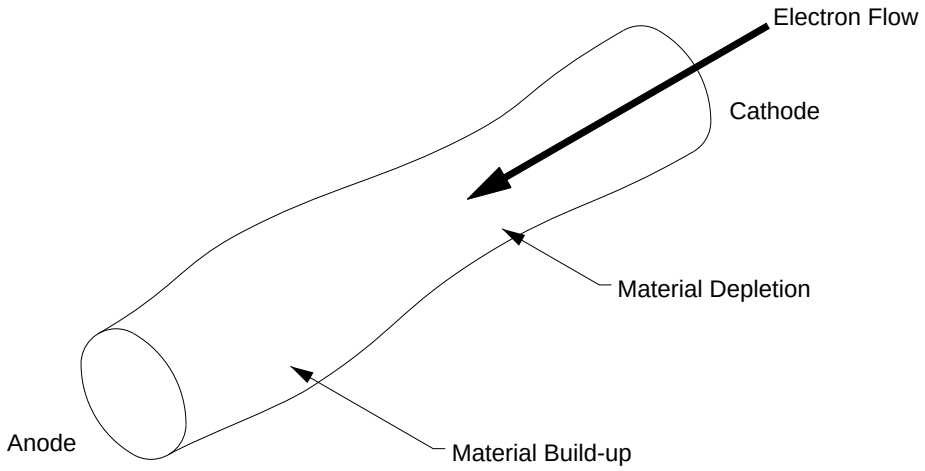

Figure 3.2: Electromigration in a wire. This figure shows how material is depleted at the cathode and deposited at the anode. The effect is exaggerated in this illustration.

Electromigration can occur in any conductor when high energy densities are present (greater than 10^5 A/cm^2 [21]). All powered metals within a semiconductor are potentially susceptible to electromigration. In particular, the areas of greatest concern are the thin-film metallic interconnects between device features, contacts and vias [18].

3.4.1.1 Physics of Failure

At the atomic level there are two competing forces operating on a conductor [28]. The first is a ‘direct’ force resulting from the electrical field. This force exerts an electrostatic pull on positively charged ion cores toward the cathode. The second force is the ‘wind’ force due to the scattering of electrons off the ions. This force acts in the opposite direction, toward the anode. At high current densities the ‘wind’ force is stronger than the ‘direct’ force so the diffusion of the ions is biased in the direction of the electron flow (anode or positive). Together these forces are referred to as the ‘electron-wind’ force.

The electromigration effects of the electron-wind depend on the material characteristics of the conductor. The activation energy for electromigration is dependent on properties such as the material type, the size and orientation of the grains, stress, heating and even the length of the conductor. For instance, even small additions of one material to a second can have a great impact on the conductor’s lifetime. As an example, pure bulk Al has an activation energy of 1.4 eV, but the addition of small amounts of Cu (0.3–5wt%Cu) can reduce this activation energy by about 0.5–0.8 eV [18].

Grain size and pattern also have a large impact on the effective activation energy of the material. For instance, thin films with a large grain size the activation energy ranges from 1–2 eV. For very fine grained samples, the activation energy may be as low as 0.4–0.6 eV. This illustrates the existence of grain boundary mass transport-induced damage. The damage is greatest at the triple points where three or more grains meet. These points act as nuclei for electromigration. A grain pattern eliminating these increases the electromigration resistance of a conductor. Triple points are eliminated as the conductor linewidth decreases to a size smaller than the grain size. At this point the line grain begins to resemble bamboo, i.e. grains lined up end to end. This is illustrated in Figure 3.3. The effect is that for aluminum interconnects, the electromigration lifetime increases as linewidths shrink below 2 μm.

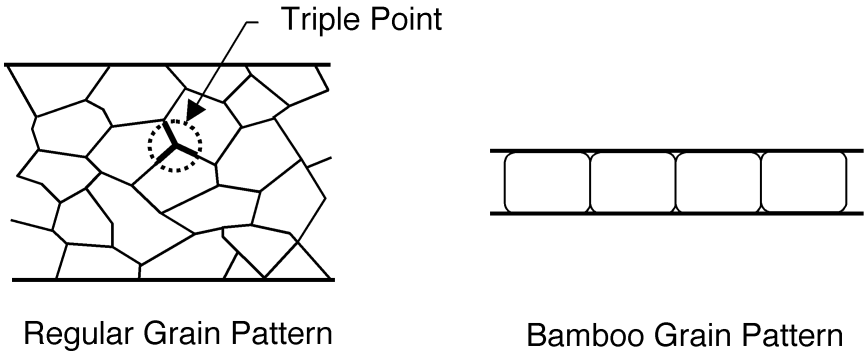

Figure 3.3: 'Bamboo' Grain Pattern. On the left is a typical grain pattern with a triple point is highlighted. On the right is a 'bamboo' pattern [21].

Another parameter affecting electromigration is stress gradients within the metal. A stress gradient can induce atomic motion within a material. Atoms migrate from regions of compressive stress to regions of tensile stress. One effect of this stress induced force is the cessation of edge migration when a conductor is shorter than a

critical length, L_c, causing the stress-induced flow of atoms to counter the electromigration movement. The result for any given current density is a critical length of conductor below which electromigration ceases [29]. This is known as the 'Blech Length'.

In addition to stress gradients, temperature gradients also have an effect on electromigration. Joule heating from high RMS currents can create thermal gradients. While these gradients may only cover a few tens of degrees temperature change, the temperature change over a few microns results in large gradients [29]. Since electromigration is a thermally activated process, the temperature gradients produce flux divergences like those found at contacts or other device features.

Increasingly, ICs have been making use of low resistivity Cu interconnects. With its lower atomic diffusivity, Cu would be expected to demonstrate a substantially improved resistance to electromigration and electromigration induced failure [30]. But this has not been the case, reliability improvements have been less than expected. The surface self-diffusion in copper appears to be faster than gain-boundary self-diffusion. Thus the surfaces provide high diffusivity paths bypassing the grain boundaries resulting in the insensitivity of copper's electromigration lifetime to different grain structures. Hau-Reige and Thompson [30] suggest that the reliability of Cu interconnects could be improved by suppressing the interface and surface diffusion. This would allow the grain structure to affect the electromigration reliability of Cu as it does in Al.

Other processes have also been applied to the fabrication of Cu interconnects, the damascene scheme in particular. An investigation by Yokogawa [31] showed that the

reliability of Cu interconnects is 50 times that of reactive ion etching (RIE) Al-Cu interconnects. He also observed single-level damascene Cu interconnects providing a 30 times longer lifetime than multi-level damascene Cu interconnects at the same current density.

3.4.1.2 Lifetime Prediction

Modeling electromigration median time to failure (MTTF) from the first principles of the failure mechanism is difficult. While there are many competing models attempting to predict time to failure from first principles, there is no universally accepted model. Currently, the favored method to predict time to failure is an approximate statistical one—*Black's* equation.

Using Black's equation, the MTTF is described by

$$MTTF = A j_e^{-n} \exp\left(\frac{E_a}{kT}\right) \tag{3.9}$$

where j_e is the current density (A/cm) and E_a is the activation energy. Failure times are described by the log-normal distribution [32]. A variation of Black's equation [22],

$$MTTF = A(j_e - j_{crit})^{-n} \exp\left(\frac{E_a}{kT}\right) \tag{3.10}$$

accounts for the Blech length. Here, j_{crit} represents the the critical current density required for electromigration with this value being inversely related to the Blech length.

Black's equation assumes activation energy is independent of line width and temperature. The symbol A is a constant dependent on a number of factors, including grain size, line structure and geometry, test conditions, current density, thermal history,

etc. Black determined the value of n to be $n = 2$. However, n is highly dependant on residual stress [25] and current density [18]. There is a great deal of disagreement on the values of n. Jensen [21] gives a range of 1–3 while Ohring [18] shows n reaching values as high as 10 when j_e approaches 1×10^7 A/cm^2. JEDEC [25] considers $n = 2$ to be valid around 1–2×10^6 A/cm^2, with the possibility of n ranging from 1–2 with variations in residual stress.

JEDEC also provides a range of values for the activation energy, E_a, of aluminum (Al) and aluminum alloys. The typical value is $E_a = 0.6$ eV with a range of 0.5–0.7 eV [25]. The activation energy can vary due to mechanical stresses caused by thermal expansion. This effect resembles a temperature dependent activation energy and can produce errors on the order of 0.1 eV. This effect is most noticeable in narrow interconnect lines under thick passivation.

Other estimates have been provided in literature. Investigation of Al-0.5% Cu interconnects [33] provided estimates of $n = 2.63$ and an activation energy of $E_a = 0.95$ eV. For multi-level Damascene Cu interconnects, the activation energy was $E_a = 0.94 \pm 0.11$ eV at a 95% confidence interval (CI) and the value of the current density exponent was found to be $n = 2.03 \pm 0.21$ (95% CI) [31].

3.4.1.3 Lifetime Distribution Model

The traditional lifetime distribution used for electromigration has been the lognormal. Most test data appears to fit well to a lognormal distribution, but this data is typically for the failure time of a single conductor [34]. Through the testing of over 75,000

Al(Cu) connectors, Gall et al [34] showed that the electromigration failure mechanism does follow the lognormal distribution. This is valid for the time to failure of the first link with the assumption that the first link failure will result in device failure. The limitation is that a lognormal distribution is not scalable. A device with different numbers of links will fail with a different lognormal distribution. Thus a measured failure distribution will only be valid for the device on which it is measured.

Additionally, in this study, Gall also showed that the Weibull (and thus the exponential) distribution is not a valid model for electromigration by demonstrating that as the number of possible failure links in a device increases, the spread of failure times decreases, meaning the β shape parameter from the Weibull distribution would have to be decreasing. If the Weibull distribution was a valid model the β would remain constant regardless of the number of possible link failures and only the characteristic life would change.

Even though the lognormal distribution is the best fit for predicting the failure of an individual device due to electromigration, the exponential model is still applicable for modeling electromigration failure in a system of many devices. This is due to the usefulness of the exponential distribution in modeling the failure rate of large systems (see Section 3.3.1).

3.4.1.4 Lifetime Sensitivity

The sensitivity of the electromigration lifetime can be observed by plotting the lifetime against varying values of the input parameters. For electromigration, the most

significant input parameters corresponding to lifetime are temperature (T) and current density (j_e). Lifetime may be non-dimensionalized by using an acceleration factor.

Substituting Black's equation (Eq. 3.9)—and assuming an exponential failure distribution—into

$$A_f = \frac{\lambda_{rated}}{\lambda} \tag{3.11}$$

provides the acceleration factor for electromigration,

$$\begin{aligned} A_{f_{EM}} &= \left(\frac{j_e}{j_{e,rated}}\right)^n \\ &\quad \times \exp\left(\left(\frac{E_{a_{EM}}}{k}\right)\left(\frac{1}{T} - \frac{1}{T_{rated}}\right)\right) \end{aligned} \tag{3.12}$$

Assuming some nominal values for E_a, j_e and T [22] we can plot the response of the acceleration factor versus scaled input parameters. Figure 3.4 shows how A_f changes for scaled values (scaling multiplier times rated value) of T and j_e ranging from 0.8–1.2 times the rated values. In this example, T has a much greater impact on A_f than j_e.

3.4.1.5 Outlook

As device features continue to shrink, and the energy densities within interconnects grow, electromigration will remain a concern. New technologies and techniques, such as Cu interconnects or designing under the Blech length, may reduce the impact of increasing densities. Historically however, as electromigration problems are eliminated, new performance demands materialize that require increased interconnect reliability under conditions where the metallization has decreased inherent reliability [29]. Because of this, electromigration will remain a design and wearout issue in future

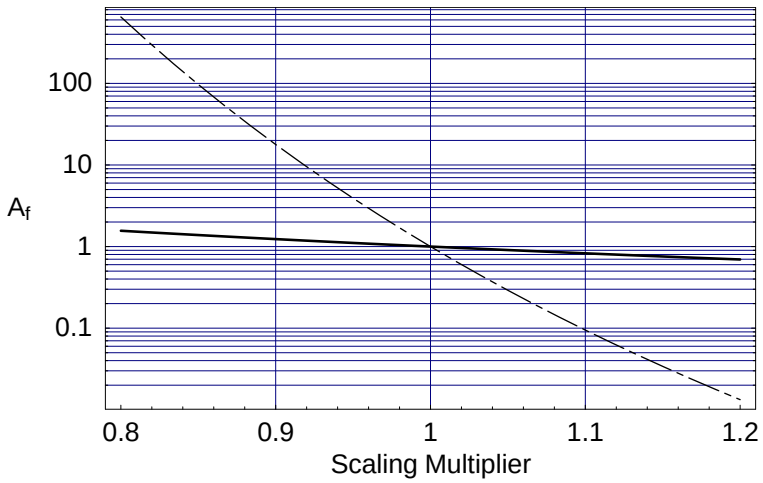

Figure 3.4: $A_{f_{EM}}$ versus scaling multiplier for T (dashed line) and j_e (solid line). $E_{a_{EM}} = 0.8\,\text{eV}$, $T = 85^{\circ}\text{C}$, $j_e = 2.5 \times 10^5\,\text{A/cm}^2$, $n = 2$.

semiconductor designs.

3.4.2 Hot Carrier Effects

Hot Carrier Effects are manifested in two distinct wearout mechanisms. These are Hot Carrier Degradation (HCD) and Negative Bias Temperature Instability (NBTI). Hot carrier effects are the result of high energy carriers, either holes or electrons, entering the gate oxide of a transistor leading to the degradation of the oxide's properties. Hot carriers are produced as current flows through the channel from the source to the drain. A small number of these hot carriers gain enough energy to be injected into the gate oxide. This results in charge trapping and the generation of interface states. Over time this leads to a shift in the performance characteristics of the device and eventually to a reduction in performance. This is referred to as HCD. Device lifetime can be

determined by defining failure as a percentage shift in threshold voltage, change in transconductance, or a variation in drive or saturation current. NBTI is caused by hole trapping and interface state generation. This results in threshold voltage shifts and delays within a CMOS device [35].

As device feature sizes continue to shrink, hot carrier effects are expected to be an increasing source of concern [36]. The rate of hot carrier degradation is directly related to the length of the channel, oxide thickness and the voltage of the device. Since the decrease of device operating voltages are chosen for optimal performance at a given life, the scaling has not kept pace with the reduction in channel length. There has been an increase in current density with a corresponding increase in semiconductor device susceptibility to hot carrier effects.

3.4.2.1 Physics of Failure

Hot carriers are generated during the operation of semiconductor devices, as it switches during a transition. As carriers travel through the channel from source to drain, the lateral electric field near the drain junction causes carriers to become hot [37]. A small percentage of these hot carriers gain sufficient energy—higher than the Si-SiO_2 energy barrier of about 3.7 eV—and the proper direction of travel to be injected into the gate oxide. In nMOS (negative-channel metal-oxide semiconductor) devices, hot electrons are generated while hot holes are produced in pMOS (positive-channel metal-oxide semiconductor) devices. Injection of either carrier results in three primary types of damage: trapping of electrons or holes in pre-existing traps, generation of new traps

and the generation of interface traps [36]. These traps may be classified by location [38] while their effects vary. Interface traps are located at or near the Si-SiO_2 interface and directly affect transconductance, leakage current and noise level. Oxide traps are located further away from the interface and affect the long term MOSFET stability, specifically threshold voltage. Effects of the defect generation includes threshold voltage shifts, transconductance degradation and drain current reduction [37]. NBTI seems to have similar degradation patterns, except for pMOS, so both will be treated as the same in this work.

Hu [39] proposed the 'lucky' electron model for hot carrier effects. This is a probalistic model built on the concept that an electron must first gain enough kinetic energy from the channel to become 'hot', and then the electron's momentum must become redirected perpendicularly so the electron can enter the oxide. The following explanation of the lucky electron model is adapted from Ohring [18].

The model begins by defining the probability that an electron can travel a distance d or more without a collision,

$$P(\text{distance} \geq d) = \mathrm{e}^{-\frac{d}{\lambda_e}} \tag{3.13}$$

where λ_e is the mean free path between scattering events, $d = \phi / q\mathscr{E}_c$, ϕ is energy, $\mathscr{E}_c$ is the channel electric field and q is the electron charge. To reach the gate oxide, electrons need to gain sufficient energy to become 'hot'. In addition, they must be redirected on a path perpendicular to the channel. The currents at the substrate (i_{sub}) and gate (i_{gate})

are indicators of the creation of hot carriers. The model for i_{sub} is

$$i_{sub} = C_1 i_{drain} \exp\left(-\frac{\beta_i}{\mathscr{E}_c}\right) \tag{3.14}$$

The exponential term comes from the dependence of the impact ionization coefficient on $1/\mathscr{E}_c$ while i_{drain} is the drain current. The parameter C_1 is a constant, or more precisely a weak function of $\mathscr{E}_c$ and the device parameters. Hu states $C_1 \approx 2$ and $\beta_i = 1.7 \times 10^6$. By defining β_i as the ratio of $\phi_i / q\lambda_e$, where ϕ_i is the minimum energy a hot electron requires in order to create impact ionization, Eq. 3.14 can be rewritten as

$$i_{sub} = C_1 i_{drain} \exp\left(\frac{-\phi_i}{\lambda_e q \mathscr{E}_c}\right) \tag{3.15}$$

Similarly the gate current is defined by

$$i_{gate} = C_2 i_{drain} \exp\left(\frac{-\phi_b}{\lambda_e q \mathscr{E}_c}\right) \tag{3.16}$$

Canceling the $\lambda_e \mathscr{E}_c$ product between Eqs. 3.15 and 3.16 results in

$$\frac{i_{gate}}{i_{drain}} = C_2 \left(\frac{i_{sub}}{C_1 i_{drain}}\right)^m \tag{3.17}$$

where $m = \phi_b / \phi_i \approx 3$ [18, 36] and represents the energy of the electrons causing damage.

During normal operation, the value of i_{gate} is negligible. Degradation due to hot-carriers is proportional to i_{gate} making gate current a good measure of the damage. If the damage resulting from HCD is designated by Δ with the time rate of change proportional to i_{gate} [18], then

$$\frac{d\Delta}{dt} \approx i_{gate} = \frac{A(\Delta)}{W} i_{drain} \left(\frac{i_{sub}}{i_{drain}}\right)^m \tag{3.18}$$

The constants C_1 and C_2 have been absorbed into $A(\Delta)$ along with an additional factor to account for the dependence of HCD on existing damage while W is the width of the MOSFET. By letting $B = A(\Delta)/W$, and knowing that MTTF depends on the reciprocal of $d\Delta/dt$, the failure rate is found from

$$\lambda = B i_{drain} \left(\frac{i_{sub}}{i_{drain}} \right)^m \tag{3.19}$$

This equation assumes static (dc) voltages and currents. To account for dynamic degradation we can use

$$\lambda = \frac{B}{T_c} \int_0^{T_c} i_{drain} \left(\frac{i_{sub}}{i_{drain}} \right)^m dt \tag{3.20}$$

where T_c is the full cycle time.

Temperature plays an interesting, though small role in hot carrier degradation. As mentioned before, the activation energy for HCD is negative, implying that HCD reduces with increasing temperature. At low temperatures, substrate current increases because drain current increases. According to Acovic [40], the effects of oxide degradation are increased at low temperatures because electrons, due to their lower thermal energy, have a hard time surmounting the potential barrier in the negatively charged degraded zone. Another possibility is that freeze out of impurities in the drain at low temperatures make nMOSFETs more sensitive to electrons trapped in the drain region, increasing degradation. Degradation decreases at high temperatures because of the decreases in drain current and mean free path.

NBTI differs from hot carrier degradation in that NBTI causes a shift in the device threshold voltage. The mechanism for NBTI damage are holes trapped within the

interface between the SiO_2 gate insulator and the Si substrate. NBTI damage is most prevalent in pMOSFET devices where holes are thermally activated and gain sufficient energy to disassociate the interface/oxide defects near the Lightly Doped Drain (LDD) regions [41]. This happens at the LDD regions because of the higher hole concentrations near the gate edge.

The first stage of the NBTI failure process begins with the generation of interface states and the production of hydrogen atoms/ions at the interface. In time, the transport of hydrogen atoms in the oxide dominates [41]. The diffusion of hydrogen is controlled by two factors. The first is the oxide field resulting from existing hole trapping and the formation of positive fixed oxide charges. The second factor is the increase in interface states. The diffusion, or generation or more interface states, is discouraged by this increase. The gradual saturation of ΔV_{th} attributed to the formation of oxide-trapped holes and fixed oxide charges which modify the oxide field to oppose the further transport of hydrogen atoms.

3.4.2.2 Lifetime Prediction

The lucky electron model does not fully answer the question on how to predict hot carrier degradation lifetime. Many other researchers have offered models for lifetime, but none are fully accepted. Since there is no direct method of measuring device lifetime [42], the Arrhenius relationship remains a favored lifetime prediction tool. The following models are from JEP-122A [22]. It contains two models, the first for nMOS device and the second for pMOS.

The *N-Channel* model is for nMOS devices. In these devices substrate current is an indicator of hot carriers. The equation is as follows:

$$MTTF = B(i_{sub})^{-N} \exp(E_a/kT) \tag{3.21}$$

where B is a scale factor as a function of doping profiles, sidewall spacing, dimensions, etc, i_{sub} is substrate current, N ranges from 2 to 4, and E_a is the activation energy in the range of -0.1 eV to -0.2 eV. In pMOS devices, hot holes do not show up as substrate current. However the gate current can serve as an indicator of hot carriers. Thus the *P-Channel* model is:

$$MTTF = B(i_{gate})^{-M} \exp(E_a/kT) \tag{3.22}$$

where B and E_a are the same as before while i_{gate} is the peak gate current during stressing and M ranges from 2 to 4. However, the Arrhenius term is not necessarily appropriate at all for these mechanisms.

A simplified version of Eq. 3.19 [36] may be used to relate lifetime prediction to the drain voltage and, ultimately, the supply voltage (V_{dd}).

$$MTTF = C \exp\left(\frac{B}{V_{dd}}\right) \tag{3.23}$$

In this equation both C and B are constants determined from life testing. The limitation is this model is valid for only a small range of gate voltages near the maximum substrate current. Here, the Arrhenius term is not used because of the small affect of temperature on HCD.

The lifetime for NBTI is described by a simple power law relationship [41]

$$MTTF = \frac{\sqrt[n]{\Delta V_{th}}}{C} \tag{3.24}$$

where ΔV_{th} is the change in threshold voltage, C is a constant, and n is the rate. As an example, in their experiment Chen et al [41] determined $n \sim 0.71$ during the reaction limited portion. This changes to approximately $n \sim 0.37$ as the process becomes diffusion controlled.

3.4.2.3 Lifetime Distribution Model

There is little discussion in literature about a proper statistical lifetime distribution model for hot carrier degradation. A logical hypothesis for the form of the lifetime distribution would be the exponential. This is a good assumption because as devices become more complex, with millions of gates, the device becomes a complex system. The probability of failure for each individual gate most likely is not an exponential distribution. But the cumulative effect of early failures and process variability, ensuring each gate has a different failure rate, widens the spread of the device failures. The end result is that the intrinsic hot carrier degradation becomes more random and statistically indistinguishable from random failures as the failures occur at a constant rate over time.

3.4.2.4 Lifetime Sensitivity

Like electromigration, the sensitivity of hot carrier degradation lifetime to changes in the input parameters may be observed. Using Eq. 3.23 as an example, the acceleration

factor for hot carrier degradation is,

$$A_{f_{HCD}} = \exp\left(B\left(\frac{1}{V_{dd}} - \frac{1}{V_{dd,max}}\right)\right) \tag{3.25}$$

The response of A_f to scaled values of V_{dd} is plotted in Figure 3.5.

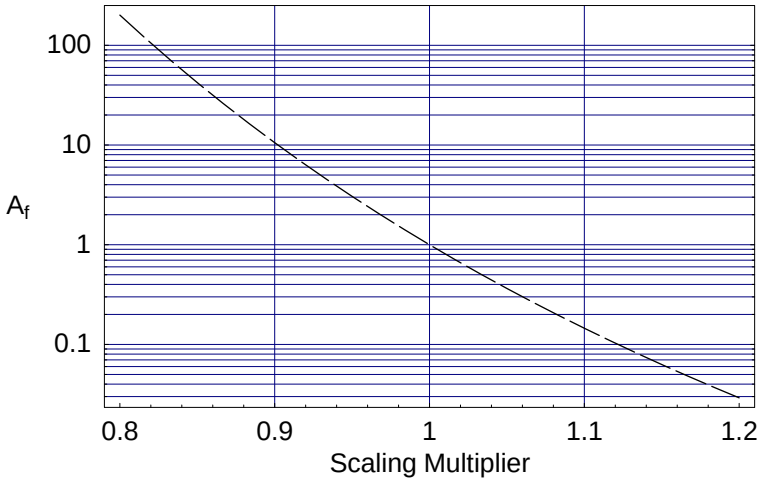

Figure 3.5: A_f versus V_{dd} scaling multiplier. $V_{dd,max} = 3.3V$ and $B = 70$ [36].

3.4.2.5 Outlook

Hot Carrier Degradation is expected to be a reliability concern as device feature sizes continue to shrink. HCD is a function of electrical fields internal to the device. Channel length, oxide thickness and device operating voltage all affect the strength of the fields and the rate of degradation. As devices are scaled downwards, channel lengths get shorter decreasing hot carrier reliability. However, the oxide thickness and voltage can also be reduced to help alleviate the reduction in reliability. Other methods of improving hot carrier reliability include possibly shifting the position of the maximum drain so it is deeper in the channel [40]. This would result in hot carriers being

generated further away from the gate and Si-SiO_2 interface, reducing the likelihood of being injected into the gate. Another method is to reduce the substrate current by using Lightly-Doped-Drain (LDD) where part of the voltage drop is a lightly doped drain extension not covered by the gate. Annealing the oxides in NH_3, N_2O or NO or growing them directly in N_2O or NO improves their resistance to interface state generation by the hot carriers.

NBTI has become a concern as device feature sizes shrunk. NBTI became evident with 0.13 μm processes as devices required much thinner gate oxides and introduced nitrides in the SiO_2 to prevent boron penetration into the gate [35]. Another source of concern is plasma-induced damage during interconnect creation resulting in driving hydrogen atoms into the Si-SiO_2 interface.

3.4.3 Time Dependent Dielectric Breakdown

Time Dependent Dielectric Breakdown (TDDB), also known as oxide breakdown, is a source of significant reliability concern for future semiconductor devices. When an electric field is applied across the dielectric gate of a transistor, the continued degradation of the material results in the formation of conductive paths and the shorting of the anode and cathode [22]. The concern is that this process will be accelerated as the thickness of the gate oxide decreases with continued device scaling.

The TDDB process takes place in two stages [43]. In the first stage, the oxide is damaged and degraded over a long time period from the localized hole and bulk

electron trapping within the oxide and at the oxide interfaces. The second stage is reached when the increasing number of traps within the oxide form a percolation (conduction) path through the oxide (see Figure 3.6). This short circuit between the substrate and gate electrode results in the failure of the oxide. This process has been successfully modeled using monte carlo simulations (a percolation model) to randomly create spherical traps.

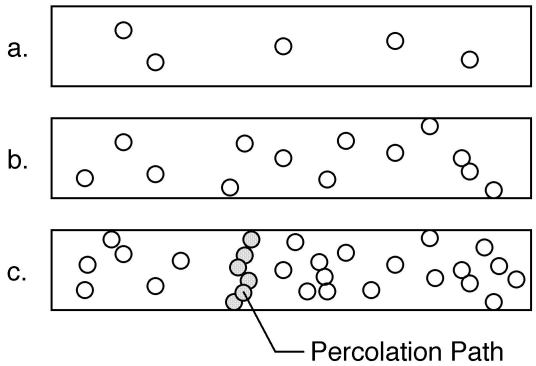

Figure 3.6: Formation of a percolation path. A small number of traps (circles) are initially in the oxide (a). Over time more traps form (b) until the number of traps is great enough to create an interconnected conduction path though the oxide (c).

The formation of the percolation path may result in one of two types of failure. When the path forms, current flow through the path causes a sudden release of energy. If this energy is sufficient it will cause runaway thermal heating and melting of the oxide, destroying the gate. This is termed *hard breakdown*. If there is not sufficient energy to result in hard breakdown, then a *soft breakdown* will result. With a soft breakdown the device continues to function. It has been speculated that soft breakdown does not even significantly affect transistor operation [44], although it may still lead to

the failure of short channel devices. However, while the change in both threshold voltage and leakage from soft breakdown is small and initially does not effect device operation, the effects are cumulative. It may be possible that multiple soft breakdowns will result in an increase in leakage current to unacceptable levels [45].

Different authors define TDDB lifetime differently. For example, one definition of lifetime is the time required for the degradation to build up to the level required for runaway [43], a hard breakdown. Alternatively, failure may be defined as the time until the first detectable electrical event [46]. This may be either a soft or hard breakdown. Given this, the definition of failure depends on the function of the device and what the TDDB effects are on that device's proper operation.

3.4.3.1 Physics of Failure

The mechanisms of oxide failure are still a subject of discussion. Several different theories to explain the phenomenon have been put forth. Two of the leading theories are the Anode Hole Injection (1/E) model and the thermochemical (E) model. Both of these models fit the experimental data in certain ranges of temperature and field, but the controversy over which is better remains. Toward this end, explanations have been published which suggest that both models are simply parts of a larger model. In contrast, the latest theories explain TDDB as having a voltage driven mechanism [47].

The Anode Hole Injection model, or 1/E model, assumes that the failure rate is inversely related to the field. Wear-out occurs as shallow traps are formed within the oxide. These traps weakly bind charge and as time progresses, new traps form and

older traps deepen. More charge is captured until a critical value Q_{BD} is reached [18]. The time to breakdown is given by

$$MTTF = \frac{Q_{BD}}{j_{FN}} = t_o \exp\left(\frac{G_R}{\mathscr{E}_{ox}}\right) \tag{3.26}$$

where j_{FN} is the tunneling current, $\mathscr{E}_{ox}$ is the oxide field, t_o is a prefactor and G_R is the field acceleration parameter. The value of G_R ranges from 290 to 350 MV/cm depending on the oxide thickness and stress type [36]. Ohring [18] defines

$$\begin{aligned} t_o &= 5.4 \times 10^{-7} \exp(-0.28\,\text{eV}/kT)\,\text{sec} \\ G_R(T) &= 120 + \frac{5.8}{kT}\,\text{MV/cm} \end{aligned}$$

where these two variables are temperature dependent and take into account hole generation and trapping efficiencies.

The thermochemical model, E-model, assumes a direct correlation between the electric field and oxide degradation. This assumption has not been conclusively proven [36], but the model does provide a good fit with experimental data . Using this model, the time to failure is given by

$$MTTF = t_o \exp(-\gamma \mathscr{E}_{ox}) \tag{3.27}$$

where t_o and γ are empirically determined constants. JEP-122A [22] defines the E-model similarly, but includes an Arrhenius term as well. Thus

$$MTTF = t_o \exp\left(\frac{E_{a_{TDDB}}}{kT} - \gamma \mathscr{E}_{ox}\right) \tag{3.28}$$

The question of which model, if either, is right is an active area of research by many parties. Both models fit the TDDB data well over limited ranges of the electric

field [48]. Groeseneken [36] states that both the E and 1/E-models are mainly valid for oxide thicknesses greater then 5 nm. Yassine [49], in testing TDDB for oxide fields ranging from 4.6 to 10.4 MV/cm, reports that the TDDB of ultrathin oxides follows the E-model down to 4.6 MV/cm, with the 1/E-model deviating from empirical data below 7.2 MV/cm. McPherson [48] argues both models are correct and are part of a complementary model where both field-induced (E-model) and current-induced (1/E-model) degradation mechanisms occur simultaneously. If one mechanism is dominant, then the model reduces to either the E or 1/E-model. For example, when electric fields are greater than 3 MV/cm (with molecular bond strengths greater than 3 eV), it reduces to the 1/E-model. When bond strengths are below 3 eV, the E-model dominates.

The debate about E vs, 1/E models is most applicable for thick oxides. For ultra-thin oxides evidence shows that gate voltage is the primary driver of the breakdown process [50]. Additionally, there is evidence that the temperature dependence of ultra-thin oxides is non-Arrhenius. Observations show the temperature acceleration factor is larger at higher temperatures . To account for these observations, Wu et al. [50] has proposed a relationship in the form of

$$MTTF = T_{BD0}(V) \exp\left(\frac{a(V)}{T} + \frac{b(V)}{T^2}\right) \tag{3.29}$$

where $T_{BD0}(V)$ is a voltage dependent prefactor and a and b are both voltage dependent as well. The second order term, b/t^2, is included in order to account for any non-Arrhenius temperature effects. Values for the terms were not yet determined.

3.4.3.2 Lifetime Distribution Model

The model typically used in accelerated testing to extrapolate failure time is the lognormal distribution. However, for TDDB the Weibull distribution provides a more accurate fit for large samples of time to failure data [46]. As oxide thickness decreases the time to failure distribution becomes wider. This results in the Weibull shape parameter decreasing as the gate oxide thickness decreases [46, 51, 52]. The decrease in β is a result of N_{BD} decreasing with the oxide thickness. When the oxide thickness equals the diameter of a defect, then only one defect is required to cause a failure. This corresponds to a β of one and happens with oxides about or below 2.2-2.7 nm [53]. In addition, the effect of variations in oxide thickness widen the failure distribution, further contributing to a reduction in the β value. Experimental results [50] show that the β factor is independent of temperature and voltage.

3.4.3.3 Lifetime Sensitivity

As an example, the derating factor for TDDB may be computed using the E-model as it is the most accepted model for thin oxide. The acceleration factor is

$$\begin{aligned} A_{f_{TDDB}} &= \exp\left(\gamma(\mathscr{E}_{ox} - \mathscr{E}_{ox,rated})\right) \\ &\times \exp\left(\frac{E_{a_{TDDB}}}{k}\left(\frac{1}{T} - \frac{1}{T_{rated}}\right)\right) \end{aligned} \tag{3.30}$$

Figure 3.7 plots the response of A_f versus the scaling multiplier for temperature and electric field (E_{ox}).

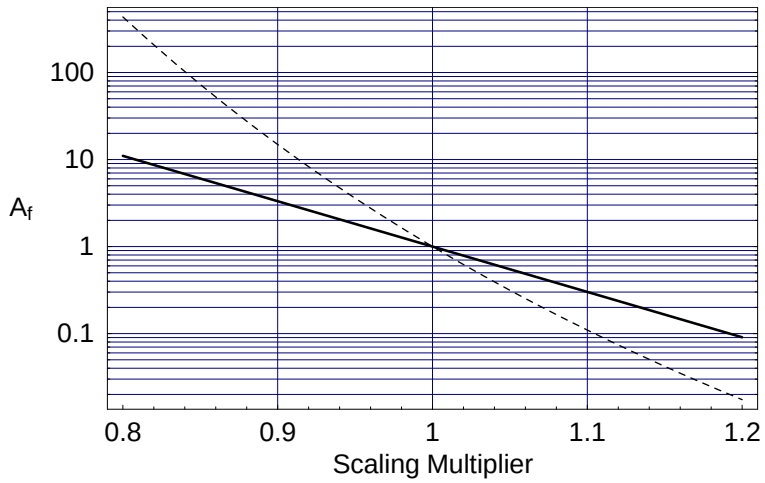

Figure 3.7: $A_{f_{TDDB}}$ versus scaling multiplier. For T (dashed line) and $\mathscr{E}_{ox}$ (solid line). $E_{a_{TDDB}} = 0.75\,\text{eV}$, $T = 85^{\circ}\text{C}$, $E_{ox} = 4\,\text{MV/cm}$, $\gamma = 3\,\text{Naperians per MV/cm}$.

3.4.3.4 Outlook

The expected high electrical fields in future semiconductor devices will have an impact on the degradation of the gate oxides. This degradation can eventually lead to a sudden breakdown of the dielectric layer and failure of the device. A complete understanding of this process does not yet exist [36] and its implications for future devices is not completely known.

Oxide thickness will continue to be scaled in future devices because of the need to improve and optimize circuit performance [43]. Groeseneken [36] shows that for an oxide thickness ranging from 4.1 nm to 1.7 nm there are several orders of magnitude drop in time to breakdown. He concludes this may be a “showstopper for the further downscaling of oxide thickness”.

In contrast to these results, Wu [54] predicts that the dielectric lifetime using the

E-model is almost infinite for a nitride/oxide in the 1.8 nm range. He further states that the existence of TDDB for ultrathin oxides under constant-current or constant-voltage stress is arguable. These contrasting viewpoints show there is still a need to better understand TDDB as device features shrink.

3.5 Impact of Technology Nodes on Lifetime

As discussed in Chapter 1, the size of semiconductor technology nodes has decreased at a steady rate and this trend is expected to continue. Substituting in the device parameters for each node into the lifetime equations from Section 3.4 allow us to see how device reliability changes due to electromigration, hot carrier and TDDB mechanisms.

The ITRS 2001 Roadmap [4] provides several tables of IC trends. Using data from the roadmap, Figure 3.8 shows how the MTTF for electromigration, TDDB and hot carrier effects vary with changing technology nodes. This plot is for high performance nMOS devices, such as microprocessors and ASICs. The vertical scale is a MTTF for each of the failure mechanisms normalized to the 2001 node which is defined as having a MTTF = 1 unit for each mode. Similar plots can be made for DRAM and low power devices. To give an idea of how years correspond to technology nodes, Table 3.1 lists the half-pitch for several years.

The graph in Figure 3.8 has some limitations which have to be taken into account when interpreting it. A major limitation is that it does not allow for any change in

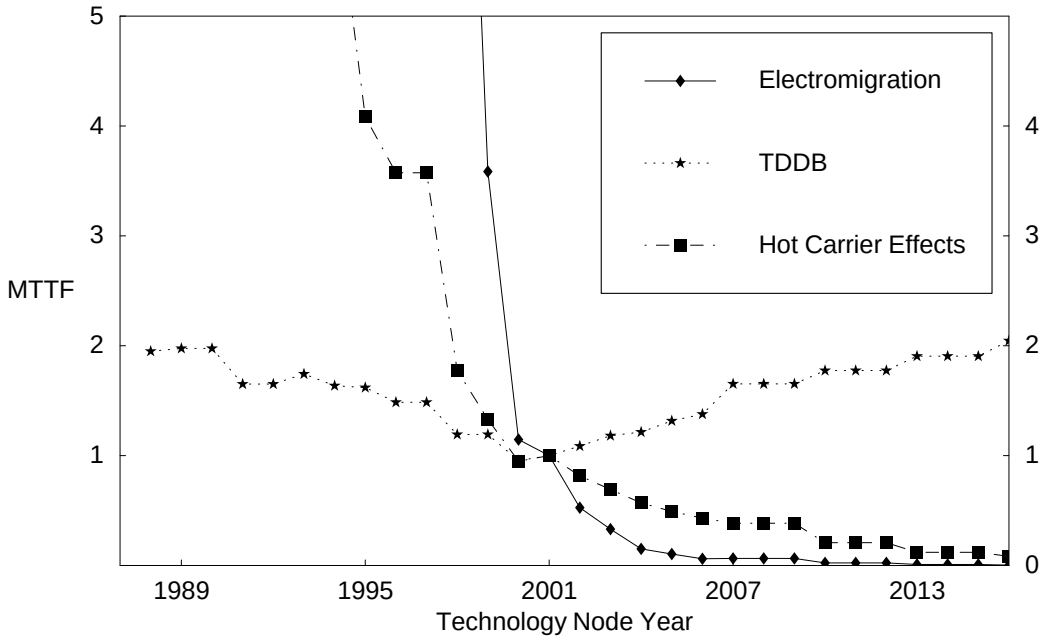

Figure 3.8: IC Reliability Trends.

Table 3.1: Technology Node Size vs. Year.

Year	Half-Pitch
2001	130 nm
2004	90 nm
2007	65 nm
2010	45 nm
2013	32 nm
2016	22 nm

technology. All the lifetime models have technology constants. Some of the change in the constants are a result of a change in manufacturing process or materials. Other changes in the technology constants result from intentional changes made to improve device reliability. As as an example of changing technology constants, consider electromigration. Figure 3.8 shows electromigration MTTF decreasing sharply. In reality electromigration reliability hasn't shown this behavior. The reason is that technology has changed. To combat the risk of electromigration, semiconductor manufacturers have made changes in the interconnect alloys used and in IC design, such as using multiple levels of metallization to reduce current densities. Additionally, every semiconductor device will inherently have different technology constants due to different design features and manufacturing processes.

The graph was built using the models presented in Section 3.4. Starting with electromigration, in Black's equation (Eq. 3.9), the paramater directly dependent on scaling is the current density, (j_e). The ITRS data provided information on feature sizes and V_{dd}, but not j_e. Using the feature size data to derive the scaling factor (κ), the change in interconnect cross-sectional area is $A_{line} = A_{line,0} 1/\kappa^2$. The current density change is $j_e \propto V_{dd}/\kappa^2$. By substituting this into Black's equation, $\text{MTTF}_{EM} \propto (V_{dd}/\kappa^2)^{-n}$. As shown in Figure 3.8, electromigration increases with scaling as the current density increases due to decreasing interconnect cross sectional area. This plot used a value of $n = 2$.

Hot carrier lifetime was determined using Eq. 3.19. Using this equation, $B = A(\Delta)/W$ where $W \propto 1/\kappa$. The value of i_{drain} is proportional to the device current

given by ITRS. This resultes in $\text{MTTF}_{HC} = i/\kappa$. For TDDB, the ITRS data directly defined the electric field across the oxide. Using this, $\text{MTTF}_{TDDB} \propto e^{\mathscr{E}}$.

As shown in Figure 3.8 and assuming unchanging technology, scaling tends to increase the failure rates for electromigration and hot carrier effects. TDDB shows an inflection around 2000 because the original ITRS data shows a corresponding peak in field strength across the dielectric at this point.

3.6 Summary

This chapter provides the background information necessary to understand the potential of device wearout and the impact of continued device scaling. The aerospace industry is concerned about the possibility of reduced semiconductor device lifetime from three major failure mechanisms, electromigration, hot carrier effects and time dependent oxide breakdown. A review of the literature shows that all three areas are valid areas of concern.

Semiconductor manufactures are aware of the problems with scaling and are working to introduce new materials and processes into their products. This is referred to as “equivalent” scaling since tradition materials are reaching their scaling limits and new materials provide a means to continue to scaling.

The semiconductor industry remains competitive and manufacturers will continue to strive for more performance and a price/performance ratio. More transistors and meters of interconnects within a IC means that the reliability per gate and unit length of

metallization must increase as the shrinking feature sizes pushes their reliability downward.

These trends will push manufactures to what Dellin [55] calls “just enough IC”. Technology will be optimized to just meet the needs of major commercial customers meaning that unnecessary reliability margins will be eliminated.

While future devices will likely meet a typical commercial customer’s reliability expectations, the reduction or elimination of reliability margins will have an impact of some non-typical customers. The military/aerospace market has effectively relied on that margin to provide the necessary reliability in COTS devices for demanding aerospace applications. Without it, their reliability needs may not be met.

As future products are developed and sold, the aerospace industry will have to remain abreast of the reliability trends. It also has to develop alternative courses of action to mitigate the impact of any loss in reliability, reduction in lifetime, or increase in wearout in these future devices.

Chapter 4

Mitigating the Impact of Decreasing Device Reliability in Aerospace Applications

The focus at the start of this research, given the concern of accelerated device wearout, was to develop design guidlines to mitigate the impact of wearout. But as discussed in chapters 1 and 3, the source of decreasing reliability is a market as well as technology driven problem. With a market driven problem, there cannot be a solely technological solution. The military/aerospace industry is a very small market for device manufactures. As such they are unlikely to devote time and money to ensure their products meet aerospace needs. There is no profit in that course of action. The aerospace industry must work within the constraint of having to use whatever devices the semiconductor industry supplies for their leading markets.

Using COTS devices in aerospace applications implies compromise. Aerospace companies gain the technological increases seen in the commercial semiconductor market and realize a reduced cost from the use of mass produced devices. Drawbacks include not having devices designed to their specific needs. After the Perry memo,

COTS devices were shown to have acceptable reliability margins for most aerospace applications [1]. With market and technology pressures reducing reliability margins, future device lifetime may be inadequate. What this means is the aerospace industry will have to adapt the available COTS devices to their needs or adapt their practices to the capabilities of the available devices. One way to accomplish this is to extend the lifetime of COTS devices. This involves reducing the stress that causes the activation of each of the potential failure mechanisms. Derating a component from its rated operating parameters is one method to reduce stress. This is a favored method of extending the lifetime of many types of mechanical and electronic components, however there is little discussion in literature of derating ICs to extend their lifetime. How this may be accomplished is the subject of the second iteration of my research. The results are presented in this chapter.

4.1 Lifetime Models—Constant Failure Rate Justification

To begin to understand how to increase lifetime, it is first necessary to understand the reliability behavior of semiconductor devices. Much of this work was comlpeted in the last chapter, but derating requires more information. Among these is a better understanding of the distribution models that may be used to represent intrinsic IC failures. This was accomplished in two ways, first by examining field failure data from avionics systems and secondly through a theoretical argument. These two approaches demonstrate it is appropriate to use a constant failure rate (exponential distribution)

model to approximate the lifetime of semiconductor devices for the purpose of predicting their lifetime improvement from derating.

The use of the exponential distribution to model the reliability of electronic components has long been a source of contention. Historically the exponential distribution has been used, and over used, to model electronic reliability. The failure rate during the ‘useful’ life portion of many electronic systems and components is often constant. Not only does the exponential distribution model this behavior, the model itself is very easy to work with and manipulate mathematically. This ease of use is so great that it results in a constant rate assumption being made even when it is inappropriate and experimental data doesn’t support its use.

4.1.1 Empirical Evidence

Analysis of empirical field data provides a glimpse of the failure behavior of in service avionics systems. Examination of existing systems, while not a direct predictor of the behavior of new systems, does provide insight into how they will behave. This section looks at two sources of empirical evidence. The first is a study of field failure data provided by members of the AVSI Project #17 team. The second set of empirical evidence comes from an older study by United Airlines on component hazard rate shapes.

4.1.1.1 Analysis of Avionics Failure Data

Two members of the AVSI Project #17 team supplied in-service field failure data. The data included the return-for-service records for eight different systems. Data from each record included serial number, date sold, date it was returned for service, replaced IC types and quantities.

The principle investigator for this analysis was Jin Qin [56]. As is typical with field data, some of the original data was incomplete or invalid. Qin reviewed the original data and discarded those records which didn't contain sufficient information to determine hours. After review, the database contained records for 18,176 systems sold between 17 August 1982 and 30 December 2001. Table 4.1 lists the size of the sample population by service year and number of failures for systems labeled A–H. Qin built the table using the following assumptions:

1. Systems were grouped by the year they entered service.
2. Records without enough information to determine service life were eliminated.
3. Only the time to first failure was calculated.
4. Censoring time of 30 April 2002.

The statistic analysis of the data was accomplished in four steps. The first step was probability plotting of the data. Qin analyzed the data using a Weibull distribution—he selected the Weibull because of its wide use in the electronics industry to represent

Table 4.1: Field Data Summary. T is the total number of system that began service in the period, F is the total number of those systems which failed [56].

	A		B		C		D		E		F		G		H	
	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F
1982											7	1				
1983			26	6							21	3				
1984			53	13							55	14				
1985			145	46							146	33			2	0
1986			64	18	310	109					185	24				
1987					433	88			58	44	283	44	32	12		
1988					798	109			149	81	428	37	116	20		
1989					665	75	4	4	296	107	429	47	279	68	3	0
1990					626	54	145	32	228	65	328	33	197	29		
1991					22	2	1303	203	183	41	431	41	312	23		
1992					8	1	1158	93	204	22	224	13	422	20	4	1
1993							661	30	190	23	125	2	376	5		
1994	38	25					586	49	122	12	25	0	365	16		
1995	80	25					423	24	77	6	32	1	298	5		
1996	106	44					494	26	97	10	39	1	268	4		
1997	220	31					671	57	158	11	29	0	452	8		
1998	233	37					730	30	153	9	26	0	221	3	285	27
1999	222	34					423	24	102	6			3	0	414	48
2000	150	16					267	9	90	7	4	0	5	0	307	6
2001	215	14					258	2	102	1	1	0			392	3

electronic component failure rates. Using a goodness-of-fit test, Qin confirmed the Weibull was indeed an appropriate lifetime model for the field data.

The second step was to estimate the Weibull parameters using the maximum likelihood estimation (MLE) technique. The critical result from this analysis was that the β parameter, the shape parameter, was close to one for each data subset.

This led to the next step, verification of the hypothesis that the data fit an exponential distribution[1]. Qin conducted a likelihood ratio test, with a significance level of 0.05, to confirm the hypothesis. The results, shown in Table 4.2, showed the exponential distribution fit most of the data subsets.

Table 4.2: Field Data Hypothesis Test. *A*: accept the hypothesis and R: reject the hypothesis [56].

	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01
A													R	A	A	A	A	A	R	A
B		A	A	A	A															
C					R	R	R	A	A	A	A									
D								A	A	A	A	A	A	A	A	A	A	A	A	A
E						A	R	R	A	R	A	R	R	A	A	R	A	A	A	A
F	A	A	A	A	R	R	R	A	A	R	A	A	A	A	A					
G						A	R	A	R	A	A	A	A	A	A	A	A			
H																	A	A	A	A

The last part of the analysis involved examining trends in the system failure rates. As shown in Figure 4.1, systems D, E, F & G show an increasing trend after 1994. System A shows this trend as well after 1998, but system H demonstrated a decreasing

[1]The exponential is a special case of the Weibull distribution where the shape parameter (β) equals one.

failure rate trend. The assumption in this analysis is that the item with a latter service entry used newer devices with smaller geometries.

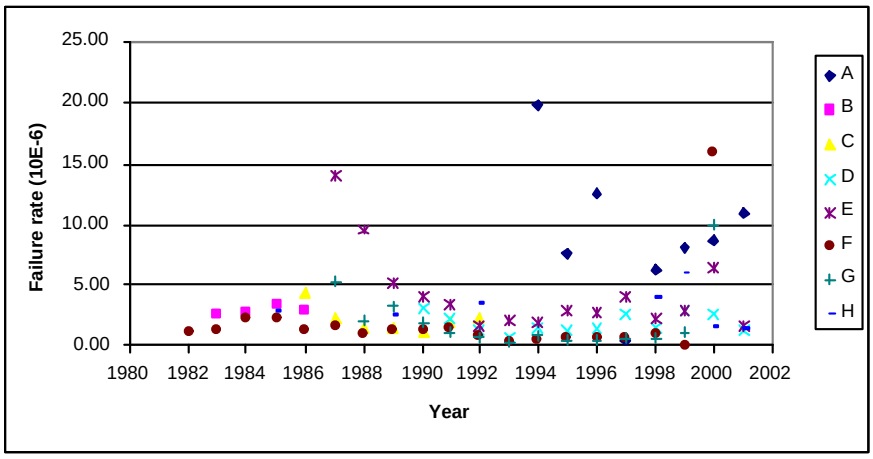

Figure 4.1: System Failure Rate Trends [56].

The conclusions drawn from this analysis were the constant failure rate model, or exponential distribution, is an appropriate model to use for inservice avionic semiconductor devices. The failure rate showed an increasing trend after 1994 confirming that the fears involved with shrinking device features are reasonable.

4.1.1.2 "Bathtub Curve Fallacy"

In the late 1960's, United Airlines examined the age related reliability patterns of the non-structural components of its aircraft fleet [57]. Using field failure rate data, they derived the hazard rate for different subsystems as a function of time. They found items that fit all types of lifetime failure rate curves including:

- Bathtub curve.
- A constant, or slowly increasing, hazard rate followed by a pronounced wearout region.
- A slowly increasing failure rate with no identifiable wearout region.
- A low initial failure probability followed by a quick rise to a constant hazard rate.
- A constant hazard rate.
- Infant mortality followed by a constant, or slowly increasing, hazard rate.

The study had several suprising conclusions. First, only six percent of aircraft components experience aging and wearout. This means most components did not follow the representational bathtub curve. A majority of components (68%), particularly electronic components, demonstrated a period of infant mortality followed by a constant, or slowly increasing, failure rate.

By itself, this study doesn't validate the use of the exponential model. The greatest shortcoming is the technology used in this study is significantly different from that found in today's, let alone future, aircraft. But the results do add support to the justification of a constant failure rate for avionics since they were conducted on electronic components operating in a similar environment.

4.1.2 Why a Constant Failure Rate Model is Justified

Empirical evidence provides a good indication that a constant rate model may be appropriate for the derating model, but this evidence is not conclusive. There are still many arguments against the exponential distribution. On the surface, one would think the exponential is not an appropriate distribution since it doesn't model either the infant mortality nor wearout phases. Intellectually, it does not seem to make sense to assume a constant failure rate since this doesn't account for aging. However, with a combination of failure mechanisms, each having a unique failure rate and distribution with a low rate of occurrence of those failures, it has become difficult to distinguise an early intrinsic wearout failure in a device from a random failure. The nature of random failures are they arrive at a constant rate.

4.1.2.1 Purpose, Scope and Assumptions

The first step in justifying the constant rate model is defining the purpose of model. All models are nothing more than approximate representations of reality. The appropriateness and usefulness of a model depends on its intended purpose and the available inputs to the model. Derived from the primary purpose of accommodating shorter lifetime devices in longlife applications, the purpose of the lifetime models for electromigration, hot carrier effects and TDDB in this report are to:

Model the change in semiconductor device lifetime, for a device operated a derated conditions, from electromigration, hot carrier and

TDDB failure mechanisms. The models shall be usable by AVSI Project #17 members, with data available to them (either from the device manufactures or via accelerated life testing) for the purpose of estimating lifetime improvement from device derating.

The scope of the model and necessary assumptions are implied by the models purpose. The scope of the derating model is to make lifetime prediction and the lifetime improvement estimates resulting from operating a semiconductor device at derated conditions.

The first assumption is it is not necessary, nor pertinent to this research, to model the failure rate over the entire lifespace of the devices. The focus of the derating model is on mean lifetime.

The second assumption is that it does not require a model with a high degree of accuracy. High accuracy would require extensive knowledge of individual IC designs, information that is proprietary to the semiconductor manufacurers and information they are unlikely to share with their aerospace customers. It would also require large and complex models which would defeat the purpose of a model that could be applied by the aerospace industry to estimate device lifetime after derating.

4.1.2.2 Failure Mechanism Lifetime Models

With the model requirements, scope and assumptions defined, the next step is to examine the individual failure mechanism lifetime models. Each of the three intrinsic failure mechanisms, electromigration, hot carrier effects and TDDB, has different

failure rate distribution as previously discussed in Section 3.4. Electromigration is generally accepted to follow a lognormal failure distribution, TDDB a Weibull distribution. Published literature on hot carrier effects had little discussion about an appropriate failure distribution for that mechanism, though the published lifetime models assumed the use of exponential [22].

Estimating TDDB lifetime using the a constant failure rate model is the easiest to justify. Several researchers have shown the Weibull shape parameter decreasing as the gate oxide thickness decreases [46, 51, 52]. Future semiconductor devices will have small dimensions with thinner oxides making the exponential model (a Weibull with $\beta = 1$) an appropriate lifetime model.

Hot carrier effects don't have an accepted lifetime distribution model. Given a lack of evidence supporting, or suggesting, any another model, the exponential is a reasonable assumption. I formal justification for this approach is the law of large numbers and complex systems demonstrating constant failure rates (see Sec. 4.1.2.3).

Electromigration is the hardest model to justify using a constraint model for since it is comonly accepted to follow the lognormal distribution. Electromigration is definitely an age related wearout mechanism with its most commonly accepted lifetime distribution being lognormal. To justify its use look back into the requirement to model the mean device lifetime. There is no need for the derating model to model the entire lifespan of the system. As device manufactures define wearout as a percentage of units failed, it is the failure during the useful life portion of the system that is of intrest.

The lognormal model has been the accepted model for electromigration lifetime

simply because it fits the data well [58]. Conceptually, the lognormal distribution has problems since it cannot be scaled with changing line length. Some authors, such as Lloyd [58] argue the Weibull distribution is a better model since electromigration failure is a 'weakest link' problem. Electromigration is considered a weakest link problem since a device's metallization is made up of a series of interconnecting links joining device features. Each link has a random inherent strength and the first to break results in failure of the chain. As the total length and quantity of interconnects increase there are more links in the chain which may break, increasing the chance of failure. The form of the cumulative distribution function (CDF) for a weakest link model is

$$F_n(s) = 1 - (1 - F_1(s))^N \tag{4.1}$$

where $F_1(s)$ is the CDF for the strength of an individual link and N is the number of individual links. The Weibull distribution models this situation opening the potential for a constant rate assumption if the β is low enough.

Other authors, such as Gall [34], have argued against using the Weibull. However, Gall's argument was that he found the Weibull slope to be decreasing as the number of possible failure links in a device increases. This argument, while increasing doubt as to the appropriateness of the Weibull, does advance the argument for justifying the use of the exponential. The number of transistors in future devices is expected to continue following Moore's Law, doubling every eighteen to twenty-four months. This means the number of links subject to electromigration doubles, pushing β lower, and making the exponential distribution an increasingly better approximation of electromigration

lifetime.

4.1.2.3 Complex System Lifetime Models

Semiconductor devices are highly complex systems with millions of individual components (As an example, the Intel Pentium IV processor has 55 million transistors [3]). Complex systems tend to exhibit constant failure rates [26]. Drenick explained this behavior within his limit theorem. He stated the reliability of a system approaches the limit given by the survival function $S(t) = e^{-\lambda t}$ as the system becomes increasingly complex. Abernathy adds further evidence when he stated that as the number of failure modes in a system increases to five or more, the Weibull shape parameter (β) will tend toward one unless all the modes have the same β and similar characteristic life.

An example of how increasing complexity results in a constant failure rate is Gall's observation of the decease in Weibull slope as the number of possible electromigration failure links in a device increases. Each of those links has a strength associated with it. That strength will vary with some distribution based on variables from design and manufacture. The stress each link will see is also a random variable, again based on the device's design and manufacture. It is possible for the strength distribution for the links to have outliers due to random non-lethal defects introduced during device manufacture. This series of random strengths, and the possibility of some lower strength links, and stresses produces a large spread in the probability distribution function (PDF) for failure of the weakest link. With enough links the PDF looks constant.

In short, complex systems fail with a constant failure rate because of the law of large numbers. In a weakest link system, failure can occur at any point as long as there is a random distribution associated with the failure time of each individual link. Even if the failure distribution of a individual link has a small variance, with a sufficiently large number of links there is a probability of a link failing at any given point. With an very large number of links the probability of failure any any given point is constant. This argument alone justifies the use of a constant failure for semiconductor devices with tens of millions of device features.

4.1.2.4 Constant Failure Rate Summary

Combining all these pieces of evidence provides a strong case for making the constant failure rate assumption. First, field data has demonstated a good fit to Weibull distributions with characteristic curves (β) of about one. An examination of the individual failure mechanisms shows them to have a lower Weibull slopes as feature sizes decrease and device complexity increases. This agrees with the notion that systems will inherently demonstrate a more constant failure rate as complexity increases. With more complex semiconductor devices it will be difficult to distinguish any inherent wearout failures from a random failure.

4.2 Lifetime Enhancement Through Derating

As shown in Section 3.2.1.1, it is possible to alter a semiconductor device's lifetime by changing its operating parameters. Two parameters a designer, using a semiconductor device, may control are junction temperature, because of heat activated mechanisms, and supply voltage. A semiconductor device's operating voltage (V_{dd}) directly affects many of its parameters. These include current density (j_e) and the electric field ($\mathscr{E}_{ox}$) across the gate dielectric. Supply voltage also has a significant effect on junction temperature (T_j).

Junction temperature (T_j) is the internal operating temperature of a device. It is dependent on the power dissipated from the device (P_D), the ambient operating temperature (T_a) and the sum of the thermal impedances between the die and ambient environment (θ_{ja}). An engineer can exercise some control over each of these factors in a system design.

The relationship for determining the junction temperature is [59]

$$T_j = \theta_{ja} P_D + T_a \tag{4.2}$$

with power dissipated determined by [60]

$$P_D = KCV_{\mathrm{dd}}^2 f + i_l V_{\mathrm{dd}} \tag{4.3}$$

where V_{dd} is the supply voltage, f is the switching frequency, K is the switching factor and C is the average node capacitance. The power dissipated is the sum of both static and dynamic power dissipation. In CMOS circuits, dynamic power is the dominate

factor, accounting for at least 90% of the power dissipation [61]. Thus a first order approximation of power disspation is

$$P_D \approx P_{dynamic} = C_{\text{eff}} V_{\text{dd}}^2 f \tag{4.4}$$

where C_{eff} combines the physical capacitance and activity (number of active nodes) to account for the average capacitance charged during each $1/f$ period.

While, this shows that V_{dd} has a direct impact on junction temperature, V_{dd} has a further impact in that frequency is proportional to it as well. In a CMOS circuit, a reduction in V_{dd} results in a near linear reduction in circuit delay. This is represented by [62]

$$f = \frac{(V_{dd} - V_{th})^2}{V_{dd}} \frac{f_{max} V_{dd,max}}{(V_{dd,max} - V_{th})^2} \tag{4.5}$$

V_{th} is the threshold voltage and f_{max} and $V_{dd,max}$ are the maximum operating frequency and voltage respectively.

To determine junction temperature in relation to the source voltage, Eqs. 4.4 and 4.5 are substituted into Eq. 4.2,

$$T_j = C_{\text{eff}} \theta_{ja} V_{\text{dd}}^2 \frac{(V_{dd} - V_{th})^2}{V_{dd}} \frac{f_{max} V_{dd,max}}{(V_{dd,max} - V_{th})^2} + T_a \tag{4.6}$$

This equation is simplified by estimating the combined value of C_{eff} and θ_{ja} after setting T_j and V_{dd} in Eq. 4.6 to $T_{j,max}$ and $V_{dd,max}$ respectively and solving for $C_{\text{eff}}\theta_{ja}$,

$$C_{\text{eff}} \theta_{ja} = -\frac{T_a - T_{j,max}}{f_{max} V_{dd,max}^2} \tag{4.7}$$

Substitute this into Eq. 4.6 results in

$$T_j = T_a - \frac{V_{dd}(V_{dd} - V_{th})^2 (T_a - T_{j,max})}{V_{dd,max}(V_{dd,max} - V_{th})^2} \tag{4.8}$$

Eq. 4.8 assumes that frequency is reduced as the supply voltage is decreased as shown in Figure 4.2. This is the first step in the determination of lifetime enhancement. The next step is to relate the other failure mechanism drivers for electromigration and TDDB to V_{dd}.

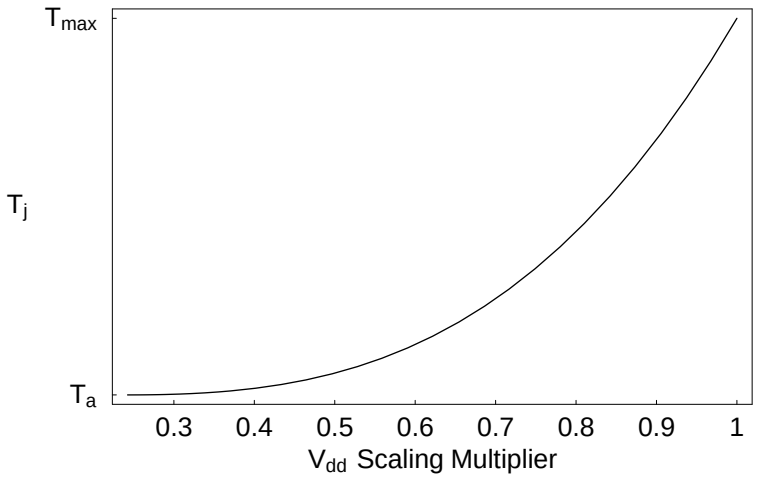

Figure 4.2: T_j given V_{dd} Scaling Multiplier

For electromigration, current density (j_e) drives wearout. It is related to V_{dd} by

$$j_e \propto \frac{V_{dd}}{A \times R} \tag{4.9}$$

where A is the cross-sectional area of the interconnect and R is the resistance. For TDDB, the oxide field is related to V_{dd} via

$$\mathscr{E}_{ox} \propto \frac{V_{dd}}{t_{\text{eff}}} \tag{4.10}$$

where t_{eff} is the effective thickness of the oxide layer.

4.2.1 Derating Factor

While the use of the term *Acceleration Factor* is appropriate in accelerated life testing, its not an intuitive term to use when discussing derating devices in order enhance their lifetime. So we will use the term *Derating Factor* (D_f). This is equivalent to A_f, but is defined as the ratio of measured MTTF of a semiconductor at its manufacturer rated operating conditions to the measured MTTF of identical devices operating at derated conditions, to more appropriately reflect its use. Symbolically this is,

$$D_f = \frac{\text{MTTF}_{\text{derated}}}{\text{MTTF}_{\text{rated}}} \tag{4.11}$$

Using this definition, our desired values for D_f are greater than zero ($D_f > 0$), with larger values providing a longer operational life. Thus derated lifetime is

$$\text{MTTF}_{\text{derated}} = D_f \times \text{MTTF}_{\text{rated}} \tag{4.12}$$

4.2.2 Modeling Voltage Derating

The effects of combined voltage and frequency derating may be seen by substituting Eq. 4.8 into the respective failure time models for each of the wearout failure mechanisms. Using the equations defined earlier, the derating factors for electromigration, hot carrier degradation and TDDB are respectively,

$$D_{f_{EM}} = \left(\frac{V_{dd,max}}{V_{dd}}\right)^n \exp\left(\left(\frac{E_{a_{EM}}}{k}\right)\left(\frac{1}{T_j} - \frac{1}{T_{j,max}}\right)\right) \tag{4.13}$$

$$D_{f_{HCD}} = \exp\left(B\left(\frac{1}{V_{dd}} - \frac{1}{V_{dd,max}}\right)\right) \tag{4.14}$$

$$D_{f_{TDDB}} = \exp\left(\gamma \mathscr{E}^{\circ}_{ox}\left(1 - \frac{V_{dd}}{V_{dd,max}}\right)\right)\exp\left(\frac{E_{a_{TDDB}}}{k}\left(\frac{1}{T_j} - \frac{1}{T_{j,max}}\right)\right) \tag{4.15}$$

For electromigration and hot carrier degradation, an advantage of having the derating factor in terms of voltage is a reduction in the number of input parameters required. For each of the wearout mechanisms, the threshold voltage (V_{th}) and rated operating voltage ($V_{dd,max}$), frequency (f_{max}) and temperature ($T_{j,max}$) are needed. For electromigration, the additional parameters required are activation energy ($E_{a_{EM}}$)) and n. The current density (j_e) is not required. For HCD, the activation energy ($E_{a_{HCD}}$)) and N are similarly required, eliminating the need for substrate current (i_{sub}). The derating factor for oxide breakdown unfortunately does not reduce the number of parameters needed, requiring the activation energy ($E_{a_{TDDB}}$)) ,γ and the oxide field ($\mathscr{E}_{ox}$).

The overall derating factor (D_f) involves combining the derating factors for each of the failure mechanisms. The derivation of combining these derating factors begins with the hazard rate. A system's hazard rate is the sum of the individual failure mode's hazard rates [20]. Assuming exponential distributions, $h_i(t) = \lambda_i = \frac{1}{\text{MTTF}_i}$. Given

$$\lambda = \sum_{i=1}^{n} \lambda_i \tag{4.16}$$

where the index i refers to each of the n failure mechanisms in turn. The failure rate of a derated IC is

$$\begin{aligned} \lambda_{derated} &= \sum_{i=1}^{n} \lambda_{derated,i} \\ &= \sum_{i=1}^{n} \frac{\lambda_i}{D_i} \end{aligned} \tag{4.17}$$

The derating factor is determined by substituting Eqs. 4.16 and 4.17 in Eq. 4.11,

$$D_f = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{n} \frac{\lambda_i}{D_i}} \tag{4.18}$$

In the case of the three wearout mechanisms discussed here

$$D_f = \frac{\lambda}{\frac{\lambda_{EM}}{D_{f_{EM}}} + \frac{\lambda_{HCD}}{D_{f_{HCD}}} + \frac{\lambda_{TDDB}}{D_{f_{TDDB}}}} \tag{4.19}$$

where λ can either represent the total failure rate or the sum of the failure rates of the wearout mechanisms. This will result in two different answers, the total derating factor and wearout derating factor respectively. Figure 4.3 shows the derating factor achieved for each of the mechanisms and the total wearout derating.

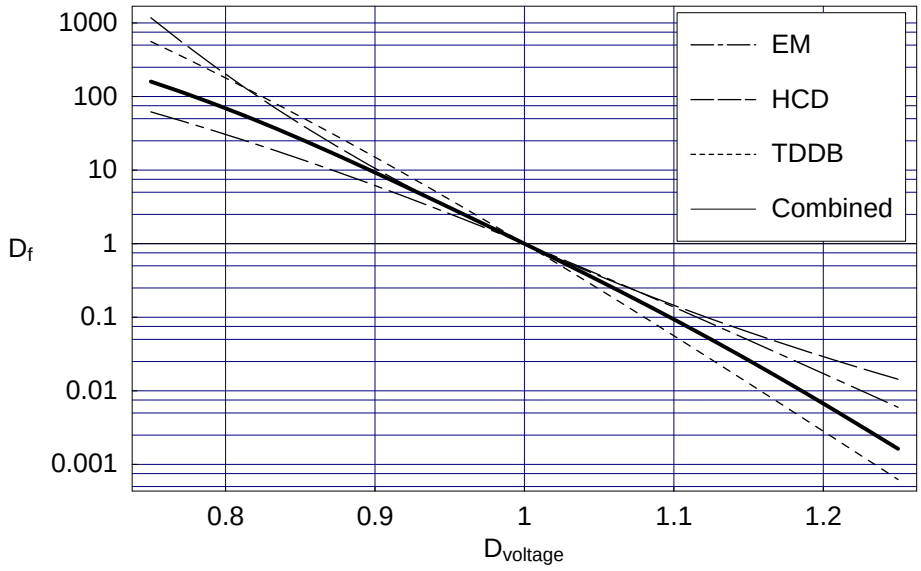

Figure 4.3: D_f versus $D_{voltage}$. $\lambda_{EM} = \lambda_{TDDB} = \lambda_{HCD}$, $T_{j,max} = 85°\text{C}$, $T_a = 20°\text{C}$, $V_{dd,max} = 3.3\,\text{V}$, $V_{th} = 0.8\,\text{V}$, $E_{a_{EM}} = 0.8\,\text{eV}$, $n = 2$, $B = 70$, $E_{a_{TDDB}} = 0.75\,\text{eV}$, $\mathscr{E}_{ox} = 4\,\text{MV/cm}$, $\gamma = 3$ Naperians per MV/cm .

To highlight the effects of the changing temperature and frequency, Figure 4.4 shows how the derating factor changes if only voltage alone is altered, leaving temperature and frequency constant. This demonstrates reducing the operating voltage alone produces some lifetime extension benefits. It also shows how a system designer

may make design trade-offs between ambient temperature, cooling and operating voltage to achieve differering levels of device performance and lifetime.

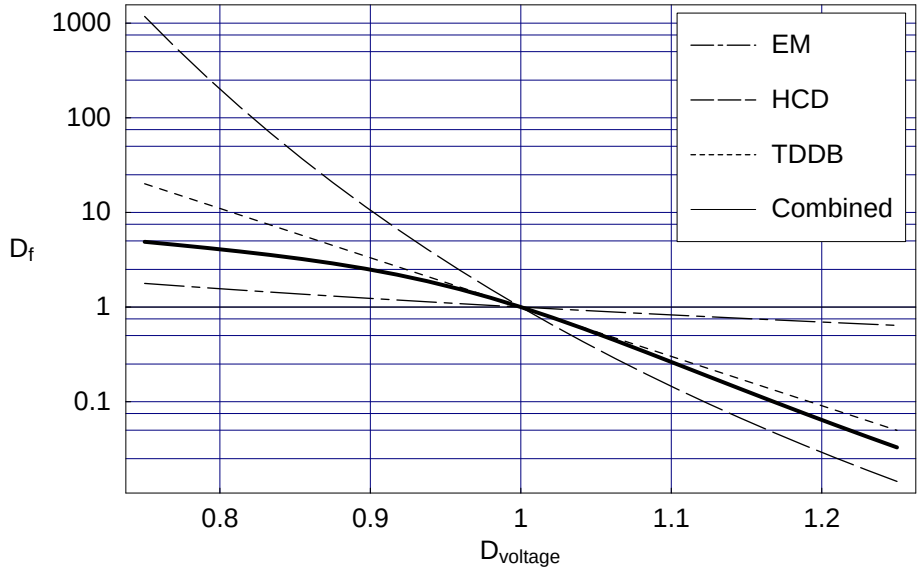

Figure 4.4: D_f versus $D_{voltage}$ with constant operating temperature and frequency. $\lambda_{EM} = \lambda_{TDDB} = \lambda_{HCD}$, $T_j = 85°\text{C}$, $T_a = 20°\text{C}$, $V_{dd,max} = 3.3\,\text{V}$, $V_{th} = 0.8\,\text{V}$, $E_{a_{EM}} = 0.8\,\text{eV}$, $n = 2$, $B = 70$, $E_{a_{TDDB}} = 0.75\,\text{eV}$, $\mathscr{E}_{ox} = 4\,\text{MV/cm}$, $\gamma = 3\,\text{Naperians per MV/cm}$.

Futhermore, the differences between figures 4.3 and 4.4 highlight a concern about applying thermal acceleration alone in accelerated life testing. Because of the low failure rates of semiconductor devices, a device's failure rate is normally determined through accelerated life testing. The failure rate of devices at accelerated conditions is determined and then extrapolate back to at-use conditions, using an acceleration factor, in order to approximate an MTTF. So when accelerated life testing is used to determine the rated lifetime of a device, care must be taken to ensure that all the relevant failure

mechanisms are accelerated in order to make a reasonable extrapolation of the device's failure rate.

4.2.3 Support for Derating

While dereating of electrical components to increase their reliability has been common practice [63], there is not much discussion of derating semiconductor devices in literature. Most likely this is because ICs have demonstrated acceptable reliability in most applications [1] and all their available performance is invariably always used. However, NASA has conducted some studies on derating for the purpose of reduced power consumption [64].

NASA conducted a study to identify flightworthy 3.3 V state-of-art semiconductor devices. The study focused on testing COTS 5 V CMOS static random access memories, from three suppliers, to see if they would function at 3.3 V. The most successful result showed one supplier's device would function over a voltage range of 3.3 V $\pm$10% and from -55 C to +125 C. This demonstates it is possible to operate devices at a lower then specified operating voltage.

In their conclusion, NASA stated many existing devices have the potential to be operated at lower than nominally specified voltages. While NASA conducted this study to validate reducing the operating voltage for reduced power consumption, this conclusion implies it will be possible to reduce operating voltage for the purpose of increasing lifetime as well.

4.3 Conclusions

The business climate is forcing the military/aerospace industry to make increased use of COTS semiconductor devices. With the strong possibility of future semiconductor devices having inadequate lifetimes for longlife aerospace applications, the aerospace industry requires a method to incorporate those devices in their systems. One way to accomplish this is through lifetime enhancement by derating the operating voltage of the devices. While derating will reduce the performance of a device, the resulting reduction of operating stresses within the device will reduce its rate of failure and extend its lifetime.

Chapter 5

Summary

5.1 Results

This research was conducted in support of AVSI Project #17, *Methods to Account for Accelerated Semiconductor Device Wearout.* The purpose of this project is understand the impact of shrinking device features, its implications on device lifetime and the potential for device wearout. This project grew out of a concern in the aerospace industry that as semiconductor device feature sizes continue to shrink, the reliability margins in new devices will no longer be sufficient to ensure they have adequate lifetime for longlife applications.

Potentially inadequate lifetime is a result of a combination of business and technological factors. The technological factors are derived from the decreasing size of device features. Smaller features, along with an increasing number of transistors, will reduce device reliability. Customers expect device manufacturers to continue to produce reliable products, even with smaller features. This is where market conditions effect the aerospace industry. Aerospace companies must use COTS devices in their

systems. As a very small portion of the semiconductor market, the aerospace industry doesn't have the purchasing clout to strongly influence device requirements. Thus when balancing reliability versus performance parameters, device manufactures will only ensure they achieve the reliability requirements of their primary customers, the computer, networking, telecommunications and consumer electronics industries. As a result, reliability of future devices is not guaranteed to meet the requirements of challenging aerospace applications.

Evidence in both published literature, and analysis of field failure data, shows there is cause to be concerned with increased failure rates from electromigration, hot carrier effects and TDDB as feature sizes decrease. A review of the failure mechanisms, lifetime and lifetime distributions for these mechanisms show the potential for increased failure rates. Combined with the business climate and market conditions, the potential for decreased device lifetime means the aerospace industry must adapt their systems to account for this limitation.

One way in which short lifetime devices my be accommodated is by extending their lifetime. This may be accomplished by derating the devices, operating them at a lower than nominal voltage. This reduced operating voltage will extend lifetime through two mechanisms. First, reducing voltage has a direct effect on the lifetime of electromigration (dependent on $J_e \propto V_{dd}$), hot carrier effects (dependent on V_{dd}) and TDDB (dependent on $\mathscr{E}_{ox}$). In addition to directly reducing lifetime, lowering the operating voltage decreases the operating temperature. This increases lifetime via the Arrhenius relationship, specifically for the electromigration and TDDB mechanisms.

The drawback to voltage derating is qualification of the devices for use and the reduction in operating frequency leading to reduced device performance.

5.2 Future Work

The work on this project remains ongoing and furtue work remains. This report presents the work accomplished to date. Three of the project Work Packages/Milestones are answered in this report. These are:

1. Determine Likely Failure Mechanisms of Future Semiconductor Devices in Avionics Applications.
2. Develop Models to Estimate Expected Lifetimes of Future Avionics.
3. Develop Device Assessment Methods and Avionics System Design Guidelines.

The remaining Work Packages/Milestones, along with what has been learned so far, serve as the basis for future work.

5.2.1 Verification and Validation of the Derating Model

The next Work Package/Milestone for the ENRE (Reliability Engineering, University of Maryland) to complete is *Verify Models*. At this point the derating model is still only theoretical. Experimental work has to be accomplished to verify and validate the model. The items that need to be verified are:

- Verifiy and Validate the derating model. Confirm the derating model makes accurate predictions of lifetime improvement for a given amount of voltage decrease. This includes verification of the constant failure rate assumption.

- Verify derated devices remain functional. The device operating characteristics and performance reduction need to be characterized.

This work has already been started at Maryland. The results of this effort will support the next future work tasks.

5.2.2 Derated Device Specification Sheets

This task will be the responsability of the AVSI member companies support by ENRE. The aerospace industry cannot derate semiconductor devices on their own. For both regulatory, business and practical reasons they are required to use devices within the data sheet specifications provided by the manufacturer. Currently, operating a device at a voltage below the specified operating limits is not supported by any known supplier [64].

To use derated COTS semiconductor devices in their systems, the aerospace industry must have the supplier characterize each derated part and publish a technical specification. The results of research so far, the derating model, and the verification of that model, provide the basis for showing that derating is a valid method of extending lifetime. Using this information, the AVSI member companies will have to build a business case for the semiconductor suppliers to develop specification sheets for

derated devices. This work may be part of the Integrated Aerospace Parts Acquisition Strategy (see Sec. 1.1.3).

5.2.3 Alternative System Architecture

There is no guarantee the aerospace industry will convince semiconductor suppliers to develop specifications for derated compinents. If implementing derated parts does not occur, aerospace companies will require an alternative method to mitigate shorter device lifetimes. One way to accomplish this is to investigate alternative avionics system design concepts. Lloyd Condra proposed the idea of a “Maintenance Free” avionics system [65]. His initial concept was:

- Develop a system architecture with small, modular, throw-away cards and standard ‘dumb’ back-planes.
- Consider each card to be a ‘system’.
- Systems can be interoperable (within and between vehicles).
- Build in redundancy and fault tolerance at the card (system) level to facilitate scheduled maintenance and eliminate unscheduled maintenance.
- Incentivize system provider with new procurement practices, e.g., power-by-the hour, etc.
- Upgrade system and insert new technology via predicted migration paths .

Ideas such as this, and other alternative concepts, need to examined to see if they would produce systems that mitigate problem of shorter lifetime devices. This work would be a much more traditional development process. It would require the development of requirements for the alternative systems, new system design concepts and the evaluation of those ideas. This task would require the use of an iterative systems engineering process to determine and refine the best system concepts.

5.3 Conclusion

In this dissertation I have validated the idea that future COTS semiconductor devices may have a shorter inherent lifetime then today's devices. The result maybe ICs having a lifetime inadequate for longlife military/aerospace applications. Electromigration, hot carrier effects and TDDB are all increased with shrinking device feature sizes, increasing the potential for failure caused by those mechanisms.

Due to market conditions and business climate, the aerospace industry must use COTS devices. While semiconductor suppliers are expected to maintain adequate device reliability for their core customers, the aerospace industry doesn't have the clout to have their requirements seriously considered. The aerospace industry must adapt COTS devices to their needs or adapt to the realities of future COTS device lifetimes. Derating an IC from its nominal operating voltage is a way to extend device lifetime and adapt COTS devices to the needs of the aerospace industry.

BIBLIOGRAPHY

[1] L. Condra, “Proposal for a contract to develop an integrated aerospace parts acquisition strategy,” July 2002.

[2] G. E. Moore, “Cramming more components onto integrated circuits,” *Electronics,* vol. 38, April 1965.

[3] “Silicon: Moore’s law.” <www.intel.com/research/silicon/mooreslaw.htm>, 2003.

[4] “International technology roadmap for semiconductors.” <public.itrs.net/Files/2001ITRS/Home.htm>, 2001.

[5] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge University Press, 1998.

[6] W. J. Perry, “Specifications and standards — a new way of doing business.” 1994.

[7] L. Condra, “Methods to account for accelerated semiconductor device wearout: AVSI project #17.” 2002.

[8] "International council on systems engineering (incose)." <www.incose.org>, 2003.

[9] M. Austin, "Lecture notes for ENSE 622/ENPM 642: Systems engineering requirements, design and trade-off analysis." March 2002.

[10] A. D. Hall, "Three-dimensional morphology of systems engineering," *IEEE transactions on Systems Science and Cybernetics*, pp. 156–160, April 1969.

[11] S. A. Fisher, D. D. Fullingim, B. L. James, J. M. Valenti, J. D. Walter, P. J. Cotter, and W. A. Seeliger, "Satellite integrated power and attitude control system design study," Master's thesis, Air Force Institute of Technology, 1997.

[12] A. P. Sage, "A case for a standard systems engineering methodology," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-7, pp. 499–504, Jul 1977.

[13] A. P. Sage, *Methodology for Large Scale Systems*. New York: McGraw Hill Book Comank, 1977.

[14] J. D. Hill and J. N. Warfield, "Unified program planning," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-2, pp. 610–621, November 1972.

[15] G. R. Mosard, "A generalized framework and methodology for systems analysis," *IEEE Transactions on Engineering Management*, vol. EM-29, pp. 81–87, August 1982.

[16] AVSI, “Project #17: Methods to account for accelerated semiconductor device wear out.” July 2001.

[17] E. A. Amerasekera, *Failure Mechanisms in Semiconductor Devices*. John Wiley & Sons, second ed., 1997.

[18] M. Ohring, *Reliability and Failure of Electronic Materials and Devices*. Academic Press, 1998.

[19] L. M. Leemis, *Reliability: Probalistic Models and Statistical Methods*. Prentice Hall, 1995.

[20] C. E. Ebeling, *An Introduction to Reliability and Maintainability Engineering*. New York: McGraw Hill, 1997.

[21] F. Jensen, *Electronic Component Reliability*. John Wiley & Sons, 1995.

[22] “Failure mechanisms and models for semiconductor devices,” tech. rep., JEDEC Solid State Technology Association, 2001.

[23] A. Mathewson, P. O’Sullivan, A. Concannon, S. Foely, S. Minehane, R. Duane, and K. Palser, “Modeling and simulation of reliability for design,” *Microelectronics Engineering*, vol. 49, pp. 95–117, 1999.

[24] W. J. Vigrass, “Calculation of semiconductor failure rates,” 2002.

[25] JEDEC, *JEP122: Failure Mechanisms and Models for Silicon Semiconductor Devices*. Electronic Industries Association, 1996.

[26] R. Riddle and J. Walter, “T-6A ARM user manual.” Air Force Operational Test & Evaluation Center, 1999.

[27] D. Young and A. Christou, “Failure mechanism models for electromigration,” *IEEE transactions on Reliability*, vol. 43, pp. 186–192, June 1994.

[28] T. O. Ogurtani, “What is electromigration?.” <www.csl.mete.metu.edu.tr/electromigration/emig.htm>, May 2001.

[29] “Electromigration for designers, an introduction for the non-specialist.” <www.cadence.com/whitepapers/whit_pdf/4095_Electromigration_WP.pdf>, 1999.

[30] C. S. Hau-Riege and C. V. Thompson, “Electromigration in Cu interconnects with very different grain structures,” *Applied Physics Letters*, vol. 78, pp. 3451–3543, 28 May 2001.

[31] S. Yokogawa, N. Okada, Y. Kakuhara, and H. Takizawa, “Electromigration performance of multi-level damascence copper interconnects,” *Microelectronics Reliability*, no. 41, pp. 1409–1416, 2001.

[32] A. H. Fischer, A. Abel, M. Lepper, A. Zitzelsbergr, and A. von Glasow, “Modeling bimodal electromigration failure distributions,” *Microelectronics Engineering*, vol. 41, pp. 445–453, 2001.

[33] B. Setlik, D. Heskett, K. Aubin, and M. A. Briere, “Electromigration investigations of aluminum alloy interconnects,”

[34] M. Gall, C. Capasso, D. Jawarani, R. Hernandez, and H. Kawasaki, “Statistical analysis of early failures in electromigration,” *Journal of Applied Physics*, vol. 90, pp. 732–740, July 2001.

[35] D. Lammers, “Reliability problems scale up as CMOS scales down,” *The Engineers Journal*, April 2002.

[36] G. Groeseneken, R. Degraeve, T. Nigam, G. V. den Bosch, and H. Maes, “Hot carrier degradation and time depedent dielectric breakdown in oxides,” *Microelectronics Engineering*, no. 49, pp. 27–40, 1999.

[37] S. Mahapatara, C. D. Parikh, V. R. Rao, C. R. Viswanathan, and J. Vasi, “Device scalling effects on hot-carrier induced interface and oxide-trappoing charge distributions in MOSFETs,” *IEEE Transactions on Electron Devices*, vol. 47, pp. 789–796, April 2000.

[38] H. Maes, G. Groesenken, R. Degraeve, J. D. Blauwe, and G. V. den Bosch, “Assessment of oxide reliability and hot carrier degradation in CMOS technology,” *Microelectronic Engineering*, no. 40, pp. 147–166, 1998.

[39] C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. W. Terrill, “Hot-carrier-induced MOSFET degradation—model, monitor, and improvement,” *IEEE Transactions on Electron Devices*, vol. ED-32, pp. 375–384, February 1985.

[40] A. Acovic, G. L. Rosa, and Y.-C. Sun, “A review of hot-carrier degradation mechanisms in MOSFETs,” *Microelectronics Reliability*, vol. 36, no. 7/8, pp. 845–869, 1996.

[41] Y. Chen, M. H. Lin, C. Chou, W. C. Wamg, S. C. Huang, Y. J. Chang, K. Y. Fu, M. T. Lee, C. H. Liu, and S. K. Fan, “Negative bias instability (NBTI) in deep sub-micro p-gate pmosfets,” in *2000 IRW Final Report*, 2000.

[42] K. Hess, L. Register, W. McMahon, B. Tuttle, O. Aktas, U. Ravaioli, J. W. Lyding, and I. C. Kizilyalli, “Theory of hot-carrier degradation in MOSFETs,” *Physica B*, vol. 272, pp. 527–531, 1999.

[43] M. T. Quddus, T. A. DeMassa, and J. J. Sanchez, “Unified model for Q_{BD} prediction for thin gate oxide MOS devices with constant voltage and current stress,” *Microelectronic Engineering*, vol. 51–52, pp. 357–372, 2000.

[44] S. Lombardo, F. Crupi, and J. H. Stathis, “Softening of breakdown in ultra-thin gate oxide nMOSFETs at low inversion layer density,” *2001 IEEE International Reliability Physics Symposium Proceedings*, vol. 39th Annual, pp. 163–167, 2001.

[45] M. Alam, B. Weir, and P. Silverman, “A future of function or failure,” *IEEE Circuits and Devices Magazine*, pp. 42–48, March 2002.

[46] J. S. Suehle, “Ultrathin gate oxide reliability: Physical models, statistics, and characterization,” *IEEE Transactions on Electron Devices*, vol. 49, pp. 958–971, June 2002.

[47] P. E. Nicollian, W. R. Hunter, and J. C. Hu, “Experimental evidence for voltage driven breakdown models in ultrathin gate oxides,” in *IEEE International Reliability Physics Symposium Proceedings*, pp. 7–15, 2000.

[48] J. W. McPherson, R. B. Khamankar, and A. Shanware, “A complemetary molecular-model (including field and current) for TDDB in SiO_2 dielectrics,” *Microelectronics Reliability*, vol. 40, pp. 1591–1597, 2000.

[49] A. Yassine, H. E. Nariman, and K. Olasupo, “Field and temperature dependence of TDDB of ultrathin gate oxide,” *IEEE Electron Device Letters*, vol. 20, pp. 390–392, August 1999.

[50] E. Wu, J. Sune, E. Nowak, J. McKenna, V. A, and D. Harmon, “Interplay of voltage and temperature acceleration of oxide breakdown for ultr-thin gate oxides,” *Solid State Electronics*, 2002.

[51] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. Maes, “A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides,” *IEDM Tech Digest*, pp. 863–866, 1995.

[52] J. H. Stahis, “Quantitative model of the thickness dependence of breakdown in ultra-thin oxides,” *Microelectronic Engineering*, no. 36, pp. 325–328, 1997.

[53] B. Weir, M. Alam, P. Silverman, F. Baumann, D.Monroe, J. Bude, G. Timp, A.Hamad, Y.Ma, M. Brown, D. Hwang, T. W. Sorsch, A. Ghetti, and G. Wilk, "Ultra-thin gate oxide reliability projections," *Solid State Electronics*, no. 46, pp. 321–328, 2002.

[54] Y. Wu, Q. Xiang, J. Y. M. Yang, G. Lucovsky, and M.-R. Lin, "Time-dependent dielectric wearout technique with temperature effect for reliability test of ultrathin (<2.0 nm) single layer and dual layer gate oxides," *Microelectronics Reliability*, vol. 40, pp. 1987–1995, 2000.

[55] T. Dellin, "Reinventing CMOS: Physics, reliability and the roadmap." March 2003.

[56] J. Qin, B. Huang, J. Walter, and J. Bernstein, "Reliability analysis of microelectronic systems in commercial aerospace industry." April 2003.

[57] M. Roush and W. Webb, *Applied Reliability Engineering II*. The Center for Reliability Engineering, University of Maryland, 2001.

[58] J. R. Lloyd and J. Kitchin, "The electromigration failure distribution: The fine-line case," *Journal of Applied Physics*, vol. 69, pp. 2117–2127, February 1991.

[59] W. H. Roadstrum and D. H. Wolaver, *Electrical Engineering For All Engineers*. Harper & Row, Publishers, Inc, 1987.

[60] J. M. Galbraith, K. F. Galloway, R. D. Schrimpf, and G. H. Johnson, “Reliability challenges for low voltage/low power integrated circuits,” *Quality and Reliability Engineering International*, vol. 12, pp. 271–279, 1996.

[61] P. Pop, “TDTS 51: Advance computer architecture lecture 11.” /webhttp://www.ida.liu.se/ TDTS51/lectures/lecture11.pdf, 2000.

[62] T. Givargis, “Uci/ics253 lecture 9.” <www.ics.uci.edu/~givargis/courses/253/notes/lecture9.pdf>, 2002.

[63] Department of Defense, *MIL-HDBK-217F*, December 1997.

[64] M. Sandor and S. Agarwal, “Low power of COTS 4M static rams for space applications,” tech. rep., NASA Jet Propulsion Laboratory, 1998.

[65] L. Condra, “Maintenance-free avionics systems: Historic challenges—historic opportunities.” January 2003.

[66] J. B. Bernstein, “Fundamentals of failure.” Class Handout, 8 October 1998.

[67] S. R. Biddle, “Reliability implications of derating leading edge high complexity microcircuits.” <www.ti.com/sc/docs/products/military/cots_pem/ridlehcm.pdf>, January 2000.

[68] A. Bravaix, D. Gouenheim, N. Revil, and E. Vincent, “Hot-carrier damage in AC-stressed deep submicrometer CMOS technologies,” in *IRW Final Report*, IEEE, 1999.

[69] T. Burd, “Low-power CMOS library design methodology,” Master’s thesis, University of California, Berkeley, 1994.

[70] P. Chaparala, J. Shibley, and P. Lin, “Thereshold drift in PMOSFETS due to NBTI and HCI,” in *IRW Final Report*, pp. 95–97, IEEE, 2000.

[71] J. A. Collins, *Failure of Materials in Mechanical Design*. New York: John Wiley & Sons, Inc., second ed., 1993.

[72] L. Condra, “Reliability assessment methods for electronic components with accelerated wearout mechanisms.” Research Proposal, May 2001.

[73] R. Degraeve, B. Kaczer, and G. Groeseneken, “Degradation and breakdown in thin oxide layers: mechanisms, models and reliability prediction,” *Microelectronics Reliability*, vol. 39, pp. 1445–1460, October 1999.

[74] R. Degreave, J. L. Ogier, R. Bellens, P. J. Roussel, G. Groeseneken, and H. E. Maes, “A new model for the field dependence of intrinsic and extrinsic time-dependent dierlectric breakdown,” *IEEE Transactions on Electron Devices*, vol. 45, pp. 472–481, February 1998.

[75] G. Delarozèe, “Introduction to reliability,” *Microelectronic Engineering*, vol. 49, pp. 3–10, 1999.

[76] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*. Duxbury Press, fourth ed., 1995.

[77] J. C. Doan, S. Lee, S.-H. Lee, P. A. Flinn, J. C. Bravman, and T. N. Marieb, "Effects of dielectric materials on electromigration failure," *Journal of Applied Physics*, vol. 89, pp. 7797–7808, June 2001.

[78] A. Dorgelo, J. Vroemen, and R. Wolters, "An additional effect of texture on the electromigration behavior of aluminum," *Microelectronic Engineering*, vol. 55, pp. 337–340, 2001.

[79] W. W. Dorner, "Using excel for weibull analysis," 1999.

[80] R. Dressen, K. Croes, J. Manca, W. D. Ceuninck, L. D. Schepper, A. Pergott, and G. Groeseneken, "A degradation model and lifetime extrapolation technique for lightly doped drain nMOSFETS under hot-carrier degradation," *Microelectronics Reliability*, vol. 41, pp. 437–443, 2001.

[81] D. D. Dylis and M. G. Priore, "A comprehensive reliability assessment tool for electronic systems," in *Annual Reliability and Maintainability Symposium*, pp. 308–313, IEEE, 22-25 January 2001.

[82] R. Gonella, "Key reliability issues for copper integration in damascene architecture," *Microelectronic Engineering*, vol. 55, pp. 245–255, 2001.

[83] F. J. Guarin, G. L. Rosa, Z. J. Yang, and S. E. R. III, "A practical approach for the accurate lifetime estimation of device degradation in deep sub-micron

CMOS technologies," in *Fourth IEEE International Caracus Conference on Devices, Circuits and Systems*, pp. D040–1–D040–8, IEEE, 2002.

[84] A. Haggag, K. Hess, W. McMahon, and L. F. Register, "Impact of scaling on CMOS IC failure rate and design rules for reliability," in *Computational Electronics, Book of Abstracts*, IWCE Glasgow 2000, 2000.

[85] C. Hu and Q. Lu, "A unified gate oxide reliability model," in *IEEE International Reliability Physics Symposium Proceedings*, pp. 47–51, IEEE, 1999.

[86] Intersil Americas, "Temperature considerations." <www.intersil.com/data/an/an9/an9207/an9207.pdf>, January 1994.

[87] *IEEE International Reliability Physics Symposium Proceedings*, 2000.

[88] J. Jones and J. Hayes, "A comparision of electronic-reliability prediction models," *IEEE Transactions on Reliability*, vol. 48, June 1999.

[89] T. Kauerauf, R. Degraeve, E. Cartier, C. Soens, and G. Groeseneken, "Low weibull slope of breakdown distributions in high-k layers," *IEEE Electron Device Letters*, vol. 23, pp. 215–217, April 2002.

[90] T. Korhonen, D. Brown, and M. Korhonen, "A grain structure based statistical simulation of electromigration damage in chip level interconnect lines," *Microelectronics Reliability*, vol. 40, pp. 2053–2060, 2000.

[91] G. L. Rosa, F. Guarin, S. Rauch, A. Acovic, J. Lukaitis, and E. Crabbe, "NBTI-Channel hot carrier effects in pMOSFETS in advanced CMOS technologies," in *IEEE International Reliability Physics Symposium*, no. 35, pp. 282 –286, 1997.

[92] D. Liu and C. Svensson, "Power consumption estimation in CMOS VLSI chips," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 6, pp. 663–670, 1994.

[93] L. D. incorperated, "Thermal considerations." <www.logicdevices.com/support/thermal.pdf>, 2002.

[94] Mathpages, "Weibull analysis," 2002.

[95] G. McFarland, *CMOS Technology Scaling and Its Impact on Cache Delay*. PhD thesis, Stanford University, June 1997.

[96] M. Modarres, M. Kaminskiy, and V. Krivtsov, *Reliability Engineering and Risk Analysis: A Practical Guide*. 270 Madision Avenue, New York, NY 10016: Marcel Dekkar, Inc., 1999.

[97] K. E. Murphy, C. M. Carter, and L. H. Wolfe, "How long should I simulate, and for how many trials? a practical guide to reliability simulations," in *Annual Reliability and Maintainability Symposium*, pp. 207–212, IEEE, 22-25 January 2001.

[98] N. D. Young and J. R. Ayres, “Negative gate bias instability in polycrystalline silicon TFT’s,” *IEEE Transactions on Electron Devices*, vol. 42, September 1995.

[99] T. H. Nig, “Silicon technology directions in the new millennium,” in *IEEE International Reliability Physics Symposium Proceedings*, pp. 1–6, 2000.

[100] M. Pecht, J. Boullie, E. Hakim, A. K. Jain, M. Jackson, I. Knowles, R. Schroeder, and A. D. Strange, “The realism of FAA reliability-safety requirements and alternatives,” *IEEE Aerospace and Electronics Systems Magazine*, vol. 13, pp. 16–20, February 1998.

[101] C. Pennetta, L. Reggiani, G. Trefán, F. Fantini, A. Scorzoni, and I. D. Munari, “A percolative approach to electromigration in metallic lines,” *Journal Of Physics D: Applied Physics*, vol. 34, pp. 1421–1429, 2001.

[102] C. Pennetta, L. Reggiani, G. Trefán, F. Fantini, I. DeMunari, and A. Scorzoni, “A percolative simulation of electromigration phenomena,” *Microelectronic Engineering*, vol. 55, pp. 349–353, 2001.

[103] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, “Impact of negative bias instability on digital circuit reliability,” in *International Reliability Physics Symposium Proceedings*, pp. 248–254, 2002.

[104] S. I. A. (SIA), “International technology roadmap for semiconductors.” <www.semichips.org>, 1999.

[105] M. Roush and W. Webb, *Applied Reliability Engineering*, vol. 1. The Center for Reliability Engineering, University of Maryland, second ed., 2000.

[106] M. Shur, *Introduction to Electronic Devices*. New York, NY: John Wiley & Sons, Inc., 1996.

[107] M. J. Skeet, "ENRE 633 final exam." 2002.

[108] J. Suné, G. Mura, and E. Miranda, "Are soft breakdown and hard breakdown of ultrathin gate oxides actually different failure mechanisms?," *IEEE Electron Device Letters*, vol. 21, pp. 167–169, April 2000.

[109] E. Takeda, E. Murakami, K. Torii, Y. Okuyama, E. Ebe, nKenji Hinode, and S. Kimura, "Reliability issues of silicon lsis facing 100-nm technology node," *Microelectronics Reliability*, no. 42, pp. 493–506, 2002.

[110] J. Walter and J. Bernstein, "The impact of semiconductor wear-out on long life applications." 2002.

[111] P.-C. Wang and R. G. Filippi, "Electromigration threshold in copper interconnects," *Applied Physics Letters*, vol. 78, pp. 3598–3600, June 2000.

[112] Y. Wu, Q. Xiang, D. Bang, G. Lucovsky, and M.-R. Lin, "Time dependant dielectric wearout (TDDW) technique for reliability of ultrathin gate oxides," *IEEE Electron Device Letters*, vol. 20, pp. 262–264, June 1999.

[113] E. Y. W. an W. W. Abadeer, L.-K. Han, S.-H. Lo, and G. R. Huckel, “Challenges for accurate reliability projections in the ultra-thin oxide regime,” in *IEEE International Reliability Physics Symposium Proceedings*, pp. 57–65, IEEE, 1999.

[114] J. F. Zhang, H. K. Sii, G. Groeseneken, and R. Degraeve, “Hole trapping and trap generation in the gate,” *IEEE Transactions on Electron Devices*, vol. 48, pp. 1127–1135, June 2001.

[115] T. Daniels, “Long life electronic systems in the new century.” May 2003.